THE UNIVERSITY
of EDINBURGH

Investigating the relationship between SORCS2 and SORLA and DNA double-stranded break formation, as a possible mechanism in neuropsychiatric disorders

Katerina Ognyanova Gospodinova

THE UNIVERSITY *of* EDINBURGH

A thesis submitted for the degree of

Doctor of Philosophy

The University of Edinburgh

2019

Declaration

I declare that this thesis has been composed by myself and that all the work presented herein is my own, except where explicitly stated otherwise by reference or acknowledgement. This work has not been submitted, in whole or in part, for any other degree or professional qualification.

Katerina Gospodinova

02/09/2019

Acknowledgements

First and foremost, I would like to thank my supervisor, Dr Kathy Evans, for being incredibly supportive, understanding and above all patient throughout the four years we have worked together. Thank you for the inspiring discussions and the complicated, 'multi-arrowed' diagrams. Thank you for giving me the freedom and space (sometimes only mental) to explore my scientific self.

Second, I would like to thank Susan Anderson without who this project and the whole institute would not have been the same. Thank you, Susan, for the enormous support and for teaching me that work is not the most important thing in life, especially after 5pm. I am also grateful to all members of the Evans group- present and past. Greatest thanks to Mairead and Rosie for the help with the statistics part of science. To Jono for the help with the 'γH2A.X staining' adventure. Many thanks also to all students I have worked with on my project.

I would also like to express my gratitude to my second supervisor- Prof Cathy Abbott, the other members of my thesis committee- Prof Tara Spires-Jones and Dr Ian Adams, as well as Prof David Porteous. Thank you for the guidance, the support and the discussions during my PhD. I am also grateful to our collaborator- Assoc. Prof Simon Glerup, and his entire group. Warmest thanks to Ditte and Sérgio, who made me feel part of the group from the first day of my visit to Denmark.

Special thanks to Shane and Fiona. I will always remember Shane's cafetière and stories about Nixon, Zizek, and many others, as well as the endless hours in TC spent with Fiona. Thank you, guys, I could not have wished for better PhD buddies.

To my Edinburgh family- Ana, Gaby and Panos. I am lacking words to express my gratitude for always being there no matter what. Thank you for the precious moments, you have made this city feel like home through the years. To Diego who has joined me in the last sprint of the PhD marathon- thank you for the great support and understanding.

Finally, to my family without whose love, support and constant encouragement I would not have done any of this. To my mother, Neli, who has been and will always be my source of inspiration to follow my instincts and wildest dreams. To my father, Ognyan, for the early lessons in (frog and worm) Biology, which later evolved into complicated discussions about the brain and the world. To my brother, Nikolay, for the wise advice in making hard decisions. To my grandparents, Stanka and Nikolay, Tzanka and Kancho, for the love and support. *Благодаря! (Thank you!)*

‘Mediocre spirits demand of science a kind of certainty which it cannot give, a sort of religious satisfaction. Only real, rare, true scientific minds can endure doubts which are attached to all our knowledge. I always envy the physicists and mathematicians who can stand on firm ground. I hover, so to speak, in thin air. Mental events seem to be immeasurable and probably always will be so.’

Sigmund Freud [Jones, 1957, Vol.2, p.418]

To them…

Abstract

SORCS2 and SORLA are members of the Vps10p-domain receptor family, which comprises five multifunctional neuronal receptors. All family members are involved in intracellular sorting and trafficking of various neurotrophic factors and their precursor forms, as well transmembrane receptors and synaptic proteins. This gene family has been implicated in a broad range of cellular processes, including neuronal health, differentiation and synaptic plasticity. Importantly for this thesis, SORLA and SORCS2 have been shown to play a role (SORLA *in vivo* and *in vitro*; SORCS2 *in vitro*) in the processing of the amyloid precursor protein (APP) and thus amyloid β (Aβ) production. Genetic and functional studies provide further evidence for the involvement of both receptors, alongside other family members, in cognition and various brain disorders.

DNA damage and compromised DNA repair have been implicated in brain aging, psychiatric and neurodegenerative disorders. Recently, a specific type of DNA damage- DNA double- stranded break (DSB) formation, has been implicated in both Alzheimer's disease (AD) and physiological brain activity, i.e. learning and memory processes. In mice, exploration of a novel environment led to a transient increase in DSB formation in the dentate gyrus (an area of the brain important for learning and memory). The breaks were repaired within 24 hours. Subsequently, it was shown that these breaks are generated by Topoisomerase II β (TopoIIβ) and are required for the expression of early-response genes. In mice, pathological Aβ levels were associated with increased neuronal DSB formation and repair deficits. The effect of Aβ on DSB formation and repair was attributed to aberrant network activity and dysregulated NMDA receptor signalling. Given the link between sortilins and neuropsychiatric illness, as well as their role in APP processing and, recently, NMDA receptor trafficking (SorCS2), my project aimed to address the relationship between the two genes and DSB formation.

I examined levels of DNA DSBs in WT and *Sorcs2*$^{-/-}$ mice before and after exploration of a novel environment. Following exploratory activity, WT mice showed increased numbers of DSBs, which were repaired within 24 hours. Meanwhile, I detected elevated levels of DSBs at baseline in the dentate gyrus of *Sorcs2*$^{-/-}$ mice. Moreover, compared to the corresponding WT mice, the *Sorcs2*$^{-/-}$ mice exhibited an altered response to the novel environment, as they failed to acquire new breaks. To explore

possible mechanisms that could explain this phenotype, I knocked out *SORCS2* in the human dopaminergic neuronal cell line LUHMES using CRISPR/Cas9 genome editing. I did not observe any difference in the extent of DSB formation in untreated *SORCS2* knockout (KO) clones, compared to WT controls grown simultaneously. However, *SORCS2* KO neurons showed elevated levels of DSBs following treatment with etoposide- a compound that prevents the re-ligation of TopoIIβ-induced DSBs. In addition, knocking out *SORCS2* was also associated with reduced neuronal viability.

I next explored possible mechanisms underlying the observed increase in etoposide induced DSBs in the *SORCS2* KO clones. In line with previous work, treatment with NMDA led to a significant increase in the number of DSBs in WT LUHMES neurons. I did not, however, observe any difference in the extent of DSB formation in NMDA-treated *SORCS2* KO clones compared to WT controls. This experiment was performed on a small number of WT and KO clones and the results may thus reflect lack of power to detect a difference. However, if this lack of difference was confirmed in subsequent experiments, this would suggest that the increase in etoposide-induced DSBs in the *SORCS2* KO clones is unlikely to be due to dysregulated NMDA signalling. Meanwhile, I was unable to measure $A\beta_{42}$ levels reliably in LUHMES. Thus, the possible role of Aβ in the increased DSB formation observed in etoposide-treated *SORCS2* KO clones cannot be excluded.

Surprisingly, mouse primary hippocampal neurons derived from *Sorcs2*$^{-/-}$ pups did not show any difference in DSB levels with or without treatment with etoposide. However, this experiment was performed on a limited number of hippocampal cultures and thus the results obtained may not reflect a true lack of difference. The possible discrepancy between the results in the etoposide-treated primary hippocampal and LUHMES dopaminergic neurons, but more importantly the role of SorCS2 in dopaminergic signalling, led me to investigate whether knocking out *SORCS2* alters extracellular dopamine levels, as a potential mechanism underlying the increased DSB formation. ELISA analysis showed no difference in dopamine release between *SORCS2* KO and WT LUHMES neurons.

Due to technical difficulties and time constraints, I was unable to assess the levels of DSB formation in *Sorl1*$^{-/-}$ mice or primary hippocampal neurons derived from them. Obtaining *SORL1* KO LUHMES cell lines using CRISPR/Cas9 genome editing also proved to be challenging. However, I generated mutant cell lines carrying the rare

SORL1 missense mutation, G508S, which is associated with early onset AD. Despite being located at a splice site, introducing the G508S mutation in LUHMES did not affect splicing. However, it led to reduced *SORL1* mRNA, but not protein, levels.

In conclusion, the work completed in this project constitutes the first evidence for a potential role of the sortilins in DSB formation. Future work investigating this link further might provide us with valuable knowledge on the cellular mechanisms underlying neurodegenerative and psychiatric conditions.

Lay Summary

Breaks in the DNA of nerve cells, or neurons, occur in the healthy brain and are normally repaired. However, if they are not repaired, these breaks can impair neuronal function and lead to premature neuronal death. Neuronal death is often seen in brain disorders, such as dementia, but its cause is unknown. The sortilins are an important family of five molecules (encoded by five genes) that help neurons function correctly and remain healthy. Mutations, or mistakes, in the sortilin genes have been discovered in a number of brain conditions. The aim of my project was to examine the link between two of the sortilin family members, SORLA and SORCS2, and the formation of DNA breaks in neurons. My experiments showed more DNA breaks in the brains of mice in which the SORCS2 molecule was missing. I also saw more breaks in human neurons that I grew in a dish after I had deleted SORCS2. Apart from having more breaks in their DNA, these neurons were also less healthy compared to neurons that had SORCS2. I was less successful in studying the link between SORLA and DNA breaks. However, I made neurons that have mutations in SORLA and can be used to study this in the future. Understanding how neuronal DNA breaks can be avoided may help us to treat brain disorders better.

Table of Contents

List of figures

List of tables

List of abbreviations

53BP1	p53-binding protein 1
8-OHdG	8-Oxo-2'-deoxyguanosine
AD	Alzheimer’s disease
ADAM10/17	A Disintegrin and metalloproteinase domain-containing protein 10/17
ADHD	Attention deficit hyperactivity disorder
AEBSF	4-(2-Aminoethyl)benzenesulfonyl fluoride hydrochloride
Akt	Protein kinase B
AP	Action potential
AP-2	Adaptor-related protein complex 2
APOE	Apolipoprotein E
APP	Amyloid precursor protein
ATM	Ataxia telangiectasia mutated
ATP	Adenosine triphosphate
ATP5B	ATP synthase F1 subunit beta
Aβ	Amyloid beta
BDNF	Brain-derived neurotrophic factor
BER	Base excision repair
bFGF	Basic fibroblast growth factor
BP	Bipolar disorder
bp	Base pair
BRCA1	Breast Cancer Gene 1
BSA	Bovine serum albumin
CA1-3	Cornu Ammonis 1-3
cAMP	Adenosine 3′,5′-cyclic monophosphate
cDNA	Complementary DNA
CHO	Chinese hamster ovary
CNS	Central nervous system
CR	Complement repeat
Ct	Cycle threshold
CTFα/β	C-terminal fragment alpha/beta
CYC1	Cytochrome C1
(d)dH_2O	(Double-)distilled water
DDR	DNA damage repair

DIV	Days *in vitro*
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNA-PK	DNA-dependent protein kinase
dNTP	Deoxynucleotide
DSB	Double-strand break
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EEF1A2	Eukaryotic Translation Elongation Factor 1 Alpha 2
EEAT3	Excitatory amino acid transporter 3
ENOX2	Ecto-NOX Disulfide-Thiol Exchanger 2
EOAD	Early-onset Alzheimer's disease
ER	Endoplasmic reticulum
Erk	Extracellular signal-regulated kinase
EST	Expressed sequence tag
FACS	Fluorescence-activated cell sorting
FBS	Fetal bovine serum
FCS	Fetal calf serum
FST	Forced swim test
FTD	Frontotemporal dementia
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GDNF	Glial cell line-derived neurotrophic factor
GFRα1	GDNF family receptor alpha-1
GGA1/2/3	Golgi-localized, gamma adaptin ear-containing, ARF-binding 1/2/3
GOI	Gene of interest
gRNA	Guide RNA
GWAS	Genome-wide association study
H2A.X	H2A histone family member X
HD	Huntington disease
HDR	Homology-directed repair
HEK293	Human embryonic kidney 293
HFS	High-frequency stimulation
HR	Homologous recombination
IF	Immunofluorescence
iPSCs	Induced pluripotent stem cells

KA	Kainic acid
KO	Knockout
LB	Lysogeny broth
LD	Linkage disequilibrium
LFS	Low-frequency stimulation
LOAD	Late-onset Alzheimer's disease
LTD	Long-term depression
LTP	Long-term potentiation
LUHMES	Lund Human Mesencephalic
MAPT	Microtubule-associated protein tau
MECP2	Methyl CpG binding protein 2
MCI	Mild cognitive impairment
MSN	Medium spiny neuron
NER	Nucleotide excision repair
NGF	Nerve growth factor
NHEJ	Non-homologous end joining
NMD	Nonsense-mediated decay
NMDAR	N-methyl-D-aspartate receptor
NPAS4	Neuronal PAS Domain Protein 4
NT	Neurotrophin
nt	Nucleotide
NTC	Non-template control
OB	Olfactory bulb
OCDL	Oxidative clustered DNA lesions
$p75^{NTR}$	Neurotrophin receptor p75
PACS1	Phosphofurin acidic cluster sorting protein 1
PAM	Protospacer adjacent motif
PAX3	Paired box gene 3
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
pHH3	Phosphorylated histone H3
PKD	Polycystic kidney disease
PLO	Poly-L-Ornithine
PSD-95	Postsynaptic density protein 95
PTZ	Pentylenetetrazol

qRT-PCR	Quantitative reverse transcription PCR
Qs	Quantity sufficient
RET	Rearranged during transfection
RMS	Rostro migratory stream
RNA	Ribonucleic acid
RNF20	Ring finger protein 20
ROS	Reactive oxygen species
RPLPO	Large Ribosomal Protein
RT	Room temperature
sAPPα/β	soluble peptide APP alpha/beta
SCLY	Selenocysteine lyase
SDHA	Succinate dehydrogenase complex flavoprotein subunit A
SDS	Sodium dodecyl sulfate
SNP	Single-nucleotide polymorphism
SORLA	Sorting-related receptor with A-type repeats
SOX2	SRY (sex determining region Y)-box 2
SSB	Single-strand break
ssODN	Single-stranded oligo DNA nucleotide
SGZ	Subgranular zone
TBE	Tris/Borate/EDTA
TBP	TATA-box binding protein
TBS	Tris-buffered saline
TGN	Trans-Golgi network
TrkA/B/C	Tropomyosin receptor kinase A/B/C
TSL	Transcript support level
TSS	Translation start site
TST	Tail suspension test
TTX	Tetrodotoxin
UBC	Ubiquitin C
UBE4	Ubiquitination factor E4
VPS10/35	Vacuolar protein sorting 10/35
VTA	Ventral tegmental area
WES	Whole-exome sequencing
WT	Wild type
XRCC4	X-Ray repair cross complementing 4

Chapter 1 Introduction

1.1. The Vps10p-domain receptor family

The Vps10p-domain receptor family, also known as the sortilins, constitutes a family of five multifunctional, type I transmembrane receptors- Sortilin in humans or sortilin in mice (encoded by the gene *SORT1* or *Sort1*, respectively), SORLA/SorLA (encoded by *SORL1/Sorl1*) and SORCS1-3/SorCS1-3 (encoded by *SORCS1-3*/*Sorcs1-3*). The receptor family is expressed predominantly in the nervous system (both developing and adult), in a dynamic, region- and cell-type specific manner. Collectively, these proteins exert diverse cellular functions, including anterograde and retrograde intracellular protein sorting and trafficking, as well as signal transduction in response to various pro- and mature neurotrophic factors. As a result, genetic variation within this receptor family has been associated with a broad range of diseases- from cardiovascular to psychiatric and neurodegenerative disorders. The role of SORLA and SORCS2 in the aetiology of neurodegenerative disorders constitutes the main focus of this thesis.

1.1.1. Structural features of the Vps10p-domain receptor family

As suggested by its name, the hallmark of the sortilin receptor family is the shared Vps10p-domain. Named after the yeast vacuolar protein sorting 10 protein (Vps10p), this is a luminal domain that is unique to the family. In yeast, Vps10p is a type I receptor with a large extracellular N-terminal domain and a short cytosolic tail, involved in the sorting of proteins to vacuoles (the yeast equivalent of lysosomes)[1].

It is considered that in mammals, the sortilins have evolved as a result of increasing cellular complexity[2] (Figure 1.1.). The Vps10p domain has been shown to harbour binding sites for a large number of ligands, including neurotrophic factors and their precursors, such as GDNF (glial cell line-derived neurotrophic factor) in the case of SorLA[3], pro- and mature NGF (nerve growth factor) and pro- and mature BDNF (brain-derived neurotrophic factor) in the case of SorCS2[4,5,6]. The role of these interactions in terms of neuronal viability and function is further discussed in the following sections.

Sortilin, which is also known as neurotensin receptor-3 (NT3), constitutes the prototype of the receptor family. The extracellular domain of sortilin consists of the

Vps10p domain, followed by a transmembrane α- helix and a short cytosolic tail[2]. In contrast, SORLA has a much more complex extracellular domain with modules not present in the rest of the family members and providing additional ligand-binding sites. For example, SORLA is the only member of the receptor family shown to bind to the amyloid precursor protein (APP), which it does via its complement-type repeat (CR) domain[7]. The exact role of the six-bladed β-propeller or the six fibronectin type-III repeats is less known, but based on ontology with other proteins they might play a role in ligand release in endocytic departments or other protein-protein interactions, respectively[8]. SORCS1-3 constitute a subfamily of homologues in which SORCS1 and SORCS3 share nearly 70% amino acid sequence identity, while SORCS2 exerts much lower sequence homology (45- 47 %) with the highest conservation found in the Vps10p domain and the lowest in the cytoplasmic tail. In the SORCS subfamily, the Vps10p domain is followed by a polycystic kidney disease (PKD) and a leucine-rich domain[2]. The exact role of the PKD domain is still unknown, but based on other proteins that share this domain, it might be involved in protein-protein interaction. While the luminal domain enables the sortilins to recognise and interact with a large number of ligands, their cytosolic tails and the recognition motifs for various adaptor proteins within them determine their role as sorting receptors (see below).

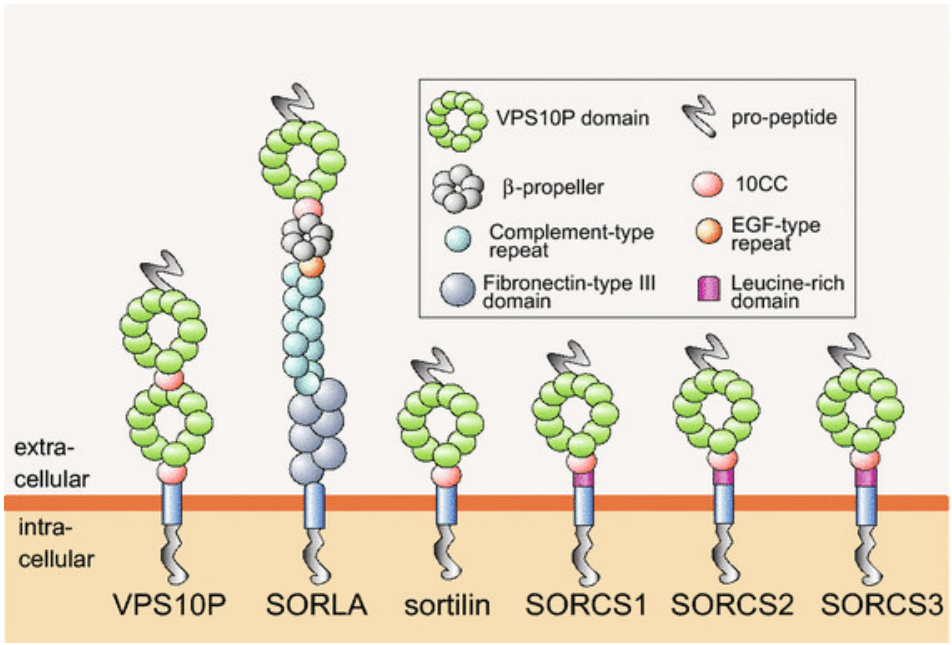

Figure 1.1. Structural organisations of the Vps10p-domain receptor family from yeast (VPS10P) and humans (SORLA, sortilin, SORCS1, -2, -3). A characteristic feature of the family is the presence of a Vps10p domain, followed by a transmembrane domain and a cytosolic tail. Additional domains are present in different members of the family. They are

mostly involved in protein-protein interactions and include leucine-rich domains (SORCS1-3), complement-type repeats, epidermal growth factor-type repeats, fibronectin-type III domains and β-propeller (SORLA). The illustration is adapted from Andersen et al. (2016)[9].

1.1.2. Processing and intracellular transport of the Vps10p-domain family

Another feature common to all members of the family is the presence of short (44-100 amino acid) propeptide at their N-terminal domain, which is proteolytically removed during the receptor's passage through the trans-Golgi network (TGN)[2] or at the cell surface[8]. In the case of SorLA and sortilin, but not SorCS1-3, this leads to receptor activation, allowing the binding of ligands to the Vps10p-domain[2]. The presence of an N-terminal propeptide has also been implicated in the processing and correct folding of newly synthesised receptors, acting as a chaperon in the endoplasmic reticulum (ER) for all family members except for SorCS2[2].

SORLA is mainly found intracellularly, with the small proportion expressed at the cell surface being rapidly internalised. The trafficking of SORLA between different cellular compartments (e.g. endosomes, lysosomes and TGN), as well as its recycling from the cell surface is mediated by adaptor complexes, which bind specific motifs within the cytoplasmic tail[10] (Figure 1.2.A). For example, SORLA internalisation at the cell surface occurs via clathrin-mediated endocytosis, which is dependent on the interaction of the receptor cytosolic tail with the clathrin adaptor protein (AP2). Once internalised, the receptor is sorted back to the TGN, where it shuttles continuously to and from endosomes. The retrograde movement of SORLA from endosomes to the TGN has been shown to be guided by the binding of PACS1 (phosphofurin acidic cluster sorting protein 1) and AP1 to the same acidic cluster in the receptor tail as AP2, as well as by the interaction of the multimeric adaptor complex retromer with another part of the cytosolic tail, known as the FANSHY motif[11,12]. Meanwhile, the receptor sorting from the TGN to endosomes depends on AP1, as well as GGA1 and GGA2- two monomeric clathrin adaptors that recognise a distinct sequence within the receptor's tail[11,13]. The ability to shuttle its cargo proteins between the TGN and endosomes constitutes one of SORLA major roles in neurons as described further in the following sections. Additionally, it has been demonstrated that this receptor might also be involved in the sorting of proteins from endosomes to both the cell surface

and lysosomes via its interaction with SN27 and GGA3, respectively[14,15]. In contrast, SorCS2 is expressed mainly at the cell surface and has been shown to contain motifs that enable its endocytosis[10].

If not internalised, all five receptors can be subjected to processing by α-secretases (ADAM10 or -17) at the cell surface, leading to shedding of their ectodomains in the extracellular space. Subsequently, what remains of the receptor is further processed by γ-secretases, releasing the receptor cytoplasmic tail[16,17] (Figure 1.2.B). The exact role of these processes and the resulting in free ectodomain and cytosolic tail remains largely unknown. However, in the case of SORLA it has been shown that once released, the receptor cytosolic tail is transported to the nucleus, where it can influence gene transcription[17].

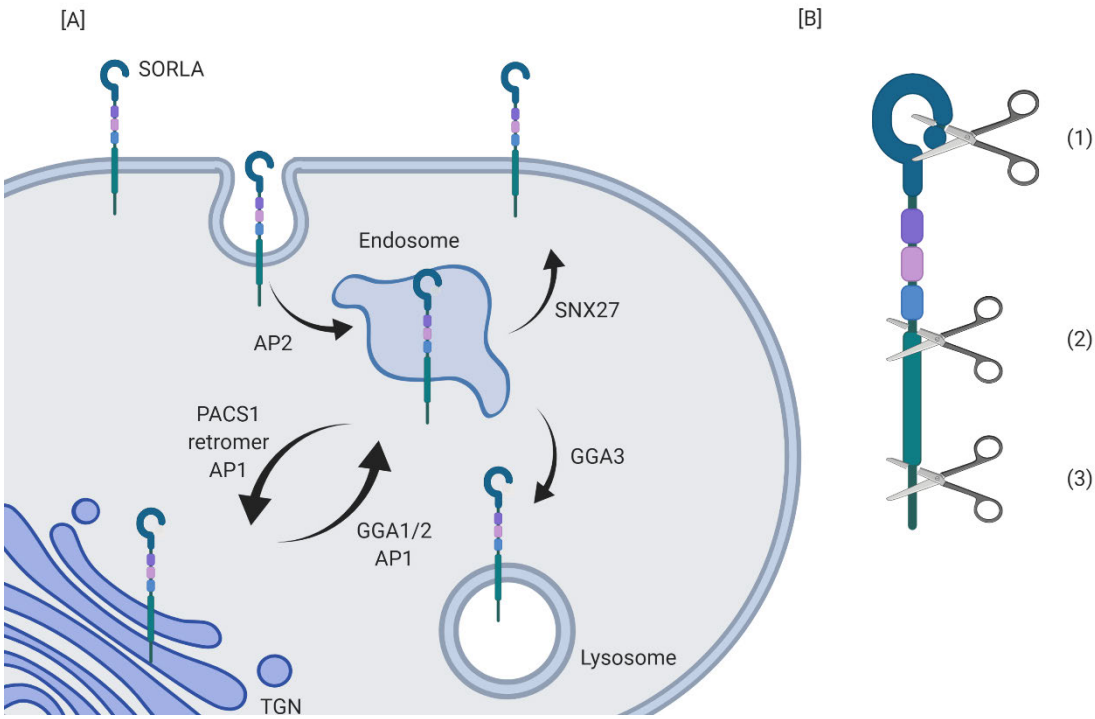

Figure 1.2. SORLA intracellular trafficking pathways and processing. [A] SORLA is synthesised as an inactive pro-receptor (pro-SORLA). Following proteolytic cleavage of the pro-peptide and maturation of the receptor in the TGN, SORLA is transported through the constitutive secretory pathway to the cell surface. This relies on the interaction of SORLA with the adaptor protein SNX27. At the cell surface, the majority of SORLA molecules undergo clathrin-mediated endocytosis facilitated by AP2. Once internalised, SORLA is constantly being recycled between the TGN and endosomal compartments. The retrograde SORLA movement from endosomes to the TGN requires SORLA interaction with the retromer complex, PACS1

and AP1. Meanwhile, the anterograde sorting between TGN and endosomes relies on SORLA interaction with GGA1 and GGA2. SORLA sorting from endosomes to lysosomal compartments is aided by the adaptor protein GGA3. [B] A figure showing the site of (1) proteolytic processing of pro-SORLA in the TGN, leading to the release of its pro-peptide and maturation of the receptor; (2) ectodomain shedding mediated by α-secretases at the cell surface and (3) release of the cytoplasmic tail following γ-secretase cleavage. [Created with BioRender]

1.1.3. Expression profile of the Vps10p-domain receptor family in the nervous system

In mice, all members of the sortilin receptor family are predominantly expressed in both the adult and developing nervous system in a dynamic and sometimes transient manner. However, members of the family have also been found in other tissues, including lung, liver, kidney and testis[1]. During embryonic development (E11.5), SorLA has been detected primarily in the cerebral cortex[18], while SorCS2 has been observed in a number of regions, including the dopaminergic midbrain nuclei and the dorsal thalamus[19]. In adult mouse and rat brains, *Sorl1* is expressed predominantly in the cortex, hippocampus and the cerebellum[18]. Meanwhile, in the mature mouse brain, *Sorcs2* has been detected in the olfactory bulb, cerebellum, amygdala, hippocampus and parts of the thalamus and midbrain[20]. While the expression patterns of *Sorl1*and *Sort1* mostly overlap, the *Sorcs* genes are expressed in a more restricted and complementary to each other manner[20]. For example, *Sorcs1* and *Sorcs3* are found primarily in the CA1 area of the hippocampus, and *Sorcs2* has been detected mainly in the CA2and dentate gyrus[21], but also in the CA1 region[5]. Moreover, it has been previously suggested that apart from being developmentally regulated, the expression of *Sorcs1* and *Sorcs3*, but not *Sorcs2*, can also be induced by neuronal activity[20]. However, a more recent paper have reported increased SORCS2 levels in the CA2 region of the hippocampus of epilepsy patients, as well as in mice upon induction of status epilepticus using chemoconvulsant agents[22]. Furthermore, *in vitro* experiments have shown *SORCS2* upregulation in an alcohol withdrawal paradigm[23]. Both epilepsy and alcohol withdrawal involve neuronal hyperexcitation.

The sortilins have been detected in defined cell populations and subcellular compartments, suggesting that they might serve distinct neuronal functions. For

example, in the adult rat brain *Sorl1* mRNA was abundant in pyramidal neurons of the frontal cortex and the CA1-3 region of the hippocampus, as well as in granule cells in the dentate gyrus and Purkinje cells in the cerebellum[18]. Meanwhile, *Sorcs2* was found in neurons in layer 5 of the cerebral cortex, pyramidal neurons of CA2 and the dentate gyrus and a subgroup of Purkinje cells in the cerebellum in the mature mouse brain[24]. At the subcellular level, SORLA has been detected mainly in the cell body, but also in dendrites, axons and axonal terminals[25,26]. Only a small proportion of these receptors (<10%) is expressed at the plasma membrane, where they are quickly internalised (see section 1.1.2.). In contrast, SorCS2 has been mainly found at the neuronal cell surface[27], and more specifically at post-synaptic densities co-localising with PSD-95 but not with the presynaptic marker synaptophysin[5].

1.1.4. Cellular roles of the Vps10p-domain receptor family in the central nervous system

The sortilins have been associated with a great number of binding partners, ranging from transmembrane receptors to soluble proteins. These interacting proteins implicate the family in multiple physiological processes related to neuronal survival, differentiation and plasticity.

1.1.4.1. The role of the Vps10p-domain receptor family in growth factor signalling

Initially, the sortilins were identified as sorting receptors involved in neurotrophic and pro-neurotrophic growth factor signalling. The neurotrophins (NTs) comprise a family of growth factors, including NGF, BDNF, NT3 and NT4/5, which are produced by neurons and activated glia cells, and are required for neuronal cell survival, growth, differentiation and synaptic potentiation. The trophic function of NTs is mediated by direct interaction with two receptor classes- the tropomyosin-related kinase receptors TrkA, -B and -C and the p75 neurotrophin receptor ($p75^{NTR}$). Co-expression with $p75^{NTR}$ strengthens Trk receptor affinity and specificity for the corresponding NTs. In contrast to the trophic role of NTs, their precursor forms (pro-NTs) have been implicated in synaptic depression, neuronal apoptosis during development, aging and neurodegeneration signalling. They also act via $p75^{NTR}$, but independently of Trk receptors[28].

Sortilin was the first member of the family to be characterised as a receptor for both pro- and mature NTs. It was shown to be involved in neuronal cell death by forming a ternary receptor complex with $p75^{NTR}$ and pro-NGF[29]. Subsequent studies in *Sort1*$^{-/-}$ mice, and neuronal cultures derived from them, further confirmed the role of sortilin in apoptotic signalling[30]. Additionally, sortilin has also been shown to interact directly with Trk receptors in hippocampal, cortical and a subtype of sensory neurons. Further experiments performed in neuronal cultures and *Sort1*$^{-/-}$ mice showed that sortilin is involved in neurotrophin signalling by facilitating anterograte axonal trafficking and synaptic localisation of Trk receptors[31]. It has also been implicated in BDNF secretion from neurons and neurosecretory cells by directly binding BDNF and mediating its intracellular trafficking to secretory granules[32].

SORLA constitutes the second member of the family to be functionally characterised. As discussed above in section 1.1.2., it interacts with many adaptor proteins involved in TGN-endosome/lysosome sorting, which enables this receptor to continuously shuttle between the plasma membrane, endosomes and the Golgi apparatus. The role of SorLA in neuronal trafficking processes was first suggested by its interaction with both GDNF and the GDNF receptors, GFRα1 (GDNF family receptor α 1) and RET (Rearranged During Transfection) receptor tyrosine kinase[3]. GDNF interacts initially with GFRα1 forming a tetrameric complex, which then binds to the RET receptor and activates multiple downstream signalling pathways, including ERK (extracellular regulated kinase) and Akt. The association of SorLA with the GDNF/GFRα1 complex leads to GDNF being targeted for lysosomal degradation, and GFRα1 being internalised from the cell surface to endosomal compartments. Moreover, SorLA also mediates the internalisation and intracellular sorting of the RET receptor to endosomes, when expressed together with GFRα1[3]. Given the importance of GDNF in the development of the dopaminergic system and the role of SorLA in the intracellular sorting of the GDNF/GFRα1/RET complex, it is not surprising that blocking the receptor activity potentiates the neurotrophic effects of GDNF in terms of cell survival, proliferation and neurite outgrowth both in SH-SY5Y cells and in primary dopaminergic neurons[3]. Additionally, knocking out *Sorl1* in mice leads to increased GDNF levels and altered dopaminergic connectivity in the striatum and midbrain[3].

SORLA has also been implicated in TrkB-dependent BDNF signal transduction[33]. Treatment of *Sorl1*$^{-/-}$ primary neuronal cultures with BDNF led to activation of its receptor TrkB, and the downstream signalling pathways ERK and Akt. However, the

strength of the induction, as assessed by the levels of TrkB, ERK and Akt phosphorylation, was much less pronounced compared to that of treated wild type (WT) neurons[33]. Subsequent experiments showed that SORLA interacts with TrkB and is required for both its anterograde and retrograde transport to and from the post-synaptic density[33]. Interestingly, BDNF itself has been previously identified as a major inducer of *Sorl1* expression, potentially leading to further potentiation of the trophic signals[34].

Similarly to sortilin, SorCS2 interacts with both pro-NGF and $p75^{NTR}$[4]. Moreover, SorCS2, but not sortilin, has been shown to be mediate pro-NGF-induced growth cone collapse in embryonic (E15) neuronal cultures grown for three days in vitro (DIV3) in the presence of $p75^{NTR}$. Further experiments demonstrated that this process required both the displacement of Trio (a scaffold protein localised at actin-rich neuronal protrusions) from the SorCS2/$p75^{NTR}$ receptor complex upon pro-NGF binding, as well as the dissociation of fascin (an actin filament stabilising protein) from the growth cone[4]. A subsequent study published in 2014 confirmed the interaction between SorCS2 and $p75^{NTR}$, both in HEK cells and brain homogenates derived from eight week-old mice[6]. Moreover, the association of these two receptors was required for the efficient binding of both pro-BDNF and pro-NGF. As previously demonstrated by Deinhardt *et al.* (2011)[4], SorCS2 co-localised with $p75^{NTR}$ in filopodia-rich protrusions and was required for proNT-induced collapse of growth cones. Neonatal cerebellar granule cells (CGN) derived from *Sorcs2*$^{-/-}$ mice failed to respond to pro-BDNF as they showed marked neurite extension compared to WT neurons, seven days post-treatment[6]. Similarly, pro-BDNF failed to induce growth cone collapse in dopaminergic neurons of midbrain explants derived from *Sorcs2*$^{-/-}$ mice (E14.5). In adult *Sorcs2*$^{-/-}$ mice, this led to increase innervation of the frontal cortex from the ventral tegmental area (VTA), suggesting that SorCS2 plays an important role in the development and maturation of the dopaminergic system through its interaction with $p75^{NTR}$ and pro-BDNF[6].

A subsequent paper showed that the interaction of SorCS2 with $p75^{NTR}$ and pro-BDNF is also required for the induction of NMDA receptor (NMDAR)-dependent long-term depression (LTD) in the hippocampus. Hippocampal slices derived from *Sorcs2*$^{-/-}$ mice were unresponsive to low-frequency stimulation (LFS) LTD protocols[5]. However, this was not the result of developmental deficits due to failure in growth cone collapse as first SorCS2 was not detected in the developing hippocampus, and second basic

synaptic properties were not altered in the knockout mice. Moreover, overexpression of *Sorcs2* in the knockout mice restored LTD, while blocking the interaction between SorCS2 and $p75^{NTR}$ in WT animals had the opposite effect[5]. Additionally, knocking out *Sorcs2* was also associated with impaired response to high frequency stimulation (HFS), known to induce NMDAR-dependent long-term potentiation (LTP). As before, lentiviral expression of *Sorcs2* in the knockout mice rescued the phenotype, while blocking SorCS2 in WT mice abolished LTP. Interestingly, high concentrations of BDNF were also sufficient to restore LTP in the *Sorcs2*$^{-/-}$ hippocampal slices[5]. Binding of BDNF to the TrkB receptor is required for the induction of the early phase of LTP, while the interaction of pro-BDNF with $p75^{NTR}$ results in LTD. Subsequent experiments showed that SorCS2 interacts with TrkB at the cell surface of hippocampal neurons and is required for the correct targeting of the receptor to post-synaptic densities upon HFS[5]. The functional importance of the interaction between SorCS2 and TrkB was further confirmed by the inability of *Sorcs2*$^{-/-}$ primary neurons to respond to BDNF treatment as demonstrated by reduced neurite branching and dendrite complexity when compared to treated WT cultures[5]. These results suggested that the interaction of SorCS2 with TrkB is crucial for both BDNF-dependent synaptic plasticity and changes in neuronal morphology[5].

1.1.4.2. The role of the Vps10p-domain receptor family in cellular processes linked to neurodegeneration

Both *in vitro* and *in vivo* functional studies (using transgenic mouse models) provide evidence for the involvement of the Vps10p-domain receptor family in neurodegeneration, and Alzheimer's disease (AD) in particular. A characteristic signature of AD pathology constitutes the abnormal deposition of 'senile plaques' and 'neurofibrillary tangles'. While AD is recognised and confirmed as a diagnosis based on the presence of these two lesions, other neuropathological abnormalities have also been observed. Substantial synaptic loss, significant cerebral cortical and hippocampal atrophy, and symmetrical enlargement of the ventricles are also strong indications of AD pathology. Nonetheless, it is worth noting that AD often occurs in conjunction with other conditions (e.g. stroke and Parkinson's disease), and when encountered in elderly, the distinction between neuropathological abnormalities occurring as a result of normal aging or dementia becomes difficult[35].

Hyper-phosphorylation of the microtubule-stabilising protein tau results in its intracellular deposition as 'neurofibrillary tangles'. Meanwhile, abnormal APP

cleavage leads to the release of insoluble Aβ, which accumulates into extracellular 'senile plaques'[35]. Newly synthesized APP molecules follow the constitutive secretory pathway, being transported from the ER through the TGN to the plasma membrane. At the plasma membrane, APP is cleaved by α-secretase, releasing the soluble APP fragment sAPPα. The membrane-anchored C-terminal fragment of APP- CTFα is internalised and further processed into peptide P3 and APP intracellular domain (AICD) by γ-secretase inside the cell. This pathway is known as the non-amyloidogenic processing pathway, as the α-secretase cleave site is within the Aβ peptide sequence, and thus Aβ oligos are not produced. In contrast, in the amyloidogenic processing pathway, unprocessed APP molecules are internalised and transported to late endosomal-lysosomal compartments, or retrogradely to the TGN. In these compartments, APP is cleaved first by β-secretase, which results in the formation of sAPPβ and the membrane-bound C terminal APP fragment- CTFβ. The latter is further processed by γ-secretases, leading to the formation of Aβ peptides of mainly 40 to 42 amino acids, and AICD[36] (Figure 1.3.A).

Functional studies have implicated SORLA in the trafficking and processing of APP, and thus Aβ accumulation. Evidence for the direct interaction between APP and SORLA was first published by Andersen *et al.* (2005)[37] who showed that SORLA can bind all human APP isoforms, forming a 1:1 stoichiometric complex. In the same study, lack of SORLA in SH-SY5Y and CHO cells that over-expressed APP, led to APP localisation mainly to the ER. Meanwhile, introducing the full length *SORL1* transcript in the same cell lines confined APP to vesicular (endosomes) and perinuclear (cis-Golgi) regions. The observed intracellular APP accumulation was accompanied by a reduction in the amount of its processing products Aβ, sAPPα and sAPPβ. In contrast, overexpression of a mutant form of *SORLA* that lacked the cytosolic domain resulted in co-localisation of APP and SORLA at the cell surface[37]. These and other studies have suggested a role of SORLA as a sorting receptor for APP, which prevents its entry into amyloidogenic and non- amyloidogenic pathways by confining the transmembrane protein to the TGN and/or recycling it between endocytic and secretory compartments. Overexpressing *SORL1* in CHO and SH-SY5Y cells prevented APP exit from the TGN, but had no effect on its trafficking from the ER to the Golgi compartments or from the cell surface to perinuclear regions[11]. This effect seemed to be dependent on SORLA interaction with the adaptor proteins PACS1 and GGA1/2. As before, introducing SORLA mutants lacking the entire cytosolic tail or carrying mutations within the PACS1 binding site led to the

mislocalisation SORLA and APP at the cell surface, accompanied by an increase in $A\beta_{40}$ and $A\beta_{42}$ levels. Meanwhile, disruption of the GGA1/2 binding sequence led to the accumulation of SORLA and APP in recycling endosomes, suggesting faulty transport to the cell surface[11]. In addition to mediating APP transport through the TGN, SORLA was also shown to influence APP processing and Aβ production by directly interacting with β- secretase in the Golgi network and thus inhibiting its access to APP[38].

In addition, further overexpression studies in HEK293 cells, followed by immunofluorescence on rat and human brain sections, showed considerable co-localization of SORLA and APP within early endosomal compartments[39]. Subsequent studies confirmed that the ability of SORLA to prevent APP processing depends on the receptor shuttling between endosome and the TGN. In a more recent study, overexpressing a mutant version of SORLA carrying a disrupted PACS1 binding site led to both the receptor and APP accumulation in endosomes[40] as opposed to the previously reported mistrafficking to the cell surface[11]. As expected, APP mislocalisation was paralleled with an increase in both amyloidogenic and non-amyloidogenic processing[40]. Similarly, introducing a version of SORLA, where the FANSHY motif has been mutated, increased APP trafficking to endosomal compartments and enhanced its amyloidogenic processing[12]. In a subsequent paper, Dumanis *et al*. (2015) replicated these observations *in vivo*. Deleting the retromer binding site in *Sorl1* resulted in the receptor depletion from the TGN and accumulation in early endosomes. Meanwhile, crossing this mouse line with the 5xFAD model of AD, which carries a total of five mutations in the human APP and presenilin-1 (*PSEN1*) transgenes, resulted in increased APP proteolytic processing, suggesting that retrograde sorting leads to the sequestration of SORLA and APP in the TGN and is required for the protection against APP amyloidogenic and non-amyloidogenic processing[15].

In mice, overexpression of human *SORL1* cDNA significantly reduced the amount of murine Aβ, but not sAPPα[41]. A similar decrease in human Aβ, but not sAPP species, was detected when human *SORL1* was expressed in PDAPP mice (an AD mouse model in which an incomplete version of the APP gene, carrying an AD-associated mutation, is expressed)[41]. The observed effect of SORLA on Aβ catabolism was shown to be the result of a direct binding of Aβ to SORLA Vps10p domain. *In vitro* experiments in SH-SY5Y cells in which WT human APP transgene was stably

expressed in the presence or absence of SORLA showed that the association between Aβ and SORLA led to increased lysosomal targeting and degradation of the Aβ peptides[41]. A subsequent study in the same cell line demonstrated that the anterograde transport of the SORLA/Aβ complex to lysosomes was mediated by the receptor interaction with the adaptor proteins GGA1 and GGA2[15]. The possible ways in which SORLA can influence APP processing and Aβ accumulation are summarised and illustrated in Figure 1.3.B.

Sorl1$^{-/-}$ mice have increased levels of Aβ, and crossing SorLA deficient mice with different AD mouse models promotes APP amyloidogenic processing even further[42,43]. In addition to the example of the double-transgenic 5xFAD mouse model described above, loss of SorLA in APP/PS1 mice (a mouse model of AD expressing the human *APP* and *PSEN1* genes bearing AD autosomal dominant mutations) accelerated the observed Aβ pathology, as well as the disease onset. Further experiments suggested that this was the result of increased APP amyloidogenic processing rather than reduced Aβ clearance[42]. Crossing *Sorl1*$^{-/-}$ mice with another mouse model of AD- AD10, in which NGF neutralisation by peripheral antibodies is associated with neurodegeneration and Aβ accumulation results in from the processing of endogenous APP, was also associated with exacerbated early Aβ pathology. Intriguingly, it also led to decreased levels of cholinergic neurodegeneration and tau hyperphosphorylation, suggesting that SorLA might be involved in various other aspects of AD-related phenotypes[43].

It is worth noting that while the above studies have provided valuable information about the mechanisms in which SORLA regulates APP processing, there are several issues surrounding the use of overexpression systems, including cytotoxicity. Moreover, the overexpression of some proteins could promote promiscuous protein-protein interactions and the abnormal modulation/activation of intracellular pathways[188]. Thus, results obtained from overexpression studies should be treated with caution, and distinction between the role of the protein of interest and the effects of the overexpression itself should be considered wherever possible.

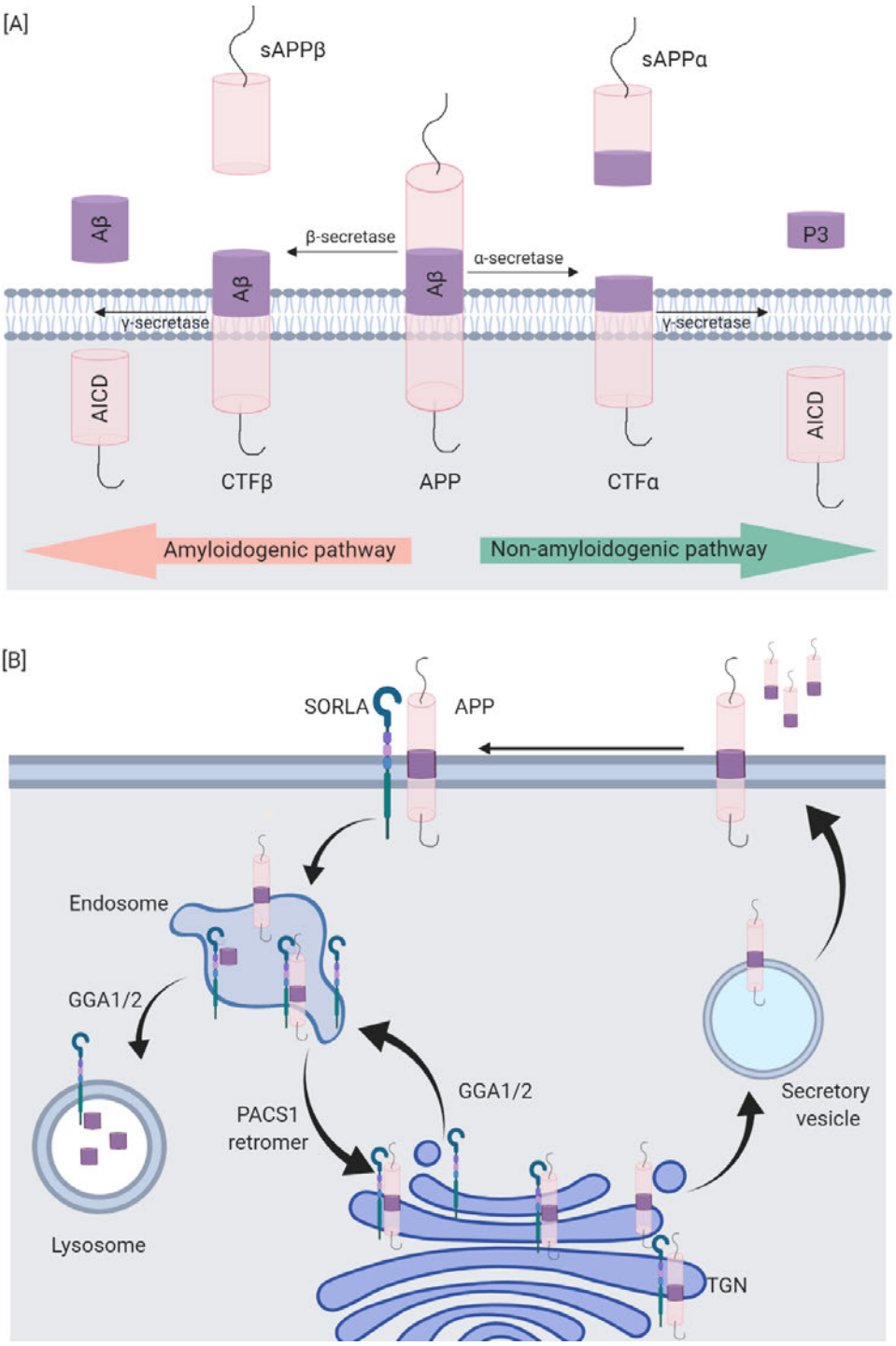

Figure 1.3. APP processing and the role of SORLA in this process. [A] APP processing pathways. In the non-amyloidogenic processing pathway, APP is initially cleaved by α-secretases at the cell surface, leading to the release of the soluble fragment sAPPα and the generation of the membrane-bound fragment CTFα. The latter is internalised and further processed into peptide P3 and AICD by γ-secretases. Alternatively, APP can be processed via the amyloidogenic pathway. Following internalisation, APP is first cleaved by β-secretases at the amino terminal end of Aβ, generating sAPPβ and CTFβ. The consequent cleavage of

CTFβ by γ-secretases at the Aβ carboxy terminal generates Aβ and AICD. [B] SORLA role in APP sorting and processing. SORLA interacts and retains APP within the TGN, hindering APP entry into amyloidogenic and non-amyloidogenic processing pathways. SORLA might also interact with APP at the cell surface, and in endosomes, following protein internalisation. Through its interaction with the retromer complex and PACS1, SORLA sorts APP from endosomal compartments to the TGN, and is recycled back through its interaction with GGA1 and GGA2. SORLA also binds Aβ directly, leading to the fragment sorting to lysosomes. This process is mediated by the interaction of SORLA with GGA1 and GGA2. [Created with BioRender].

In addition to its binding to APP and Aβ, SORLA has been implicated in the aetiology of AD through its interaction with the apolipoprotein E (APOE). APOE is a glycoprotein, mediating the transport and delivery of cholesterol and other lipids. It constitutes the major genetic risk factor for developing both EOAD and LOAD. In humans, there are three *APOE* alleles: ε3 constitutes the common variant, while ε4 increases and ε2 decreases the risk of AD[44]. *In vitro*, SORLA has been shown to bind to the APOE isoforms with differing affinity- the highest affinity was detected upon binding to ε4, followed by ε3 and finally ε2. This difference in binding corresponded to isoform-dependent cellular uptake of both APOE and APOE-Aβ complexes in the presence of SORLA (the highest uptake was observed in presence of the ε4 allele, followed by ε3 and ε2)[45].

Members of the SORCS subfamily have also been associated with APP processing and thus potentially with AD. Overexpression of all members of the family in HEK293 cells led to a significant decrease in the levels of APP processing mediated by γ-secretase. Meanwhile, knocking down *SORCS2*, *SORCS3* and *SORT1* had the opposite effect on γ-secretase activity[46]. Similarly, overexpression of both *SORCS1* and APP in HEK293 was associated with decreased secretion of $A\beta_{40}$ and $A\beta_{42}$, as well as sAPPα[47]. Meanwhile, knocking down *SORCS1* with shRNAs had the opposite effect in terms of APP processing. Immunoprecipitation experiments in HEK293 cells, supported by confocal analysis in SH-SY5Y cells, showed an interaction between SORCS1 and APP[47].

Sortilin has also been shown to interact with both APP and APOE. However, unlike SORLA, crossing sortilin knockout mice with commonly used AD mouse models (namely, PDAPP and 5xFAD mice) did not alter murine or human APP and sAPP

levels, but only increased Aβ production. *In vitro* experiments attributed the observed increase in Aβ levels to impaired cellular uptake of APOE. Thus, sortilin was identified as a major cell surface receptor for APOE, mediating the endocytosis and subsequent degradation of Aβ bound to APOE[48]. With respect to APP, sortilin has been shown to associate with the transmembrane receptor in neuronal processes. Additionally, sortilin deficiency did not affect APP intracellular distribution, but led to decreased sAPPα extracellular levels[49].

Apart from AD, almost all members of the sortilin family, except for SORCS1, have also been associated with the neurobiology of other neurodegenerative conditions. For example, functional studies have linked both SorLA and SorCS2 to Huntington's disease (HD). The striatum constitutes the most affected part of the brain in HD, and reduced BDNF levels in this area have been detected in HD patients, as well as in a mouse model of HD- the HD82 mice[50]. In the latter, the decrease in BDNF levels coincided with decrease *Sorl1* expression[34]. Based on this observation, as well as on the role of SorLA in the trafficking of the BDNF receptor TrkB, Rohe *et al.* (2013)[33] hypothesised that loss of SorLA would lead to potentiation of the BDNF-dependent neurodegenerative phenotype observed in the HD82 mice. Indeed, crossing *Sorl1*$^{-/-}$ and HD82 mice accelerated the onset of the motor deficits normally observed in these mice. In the double transgenic mice, hind limb clasping and compromised rotarod performance appeared at around ten weeks of age in contrast to the HD82 mice, which normally displayed these phenotypes between 16-22 weeks of age[33]. Similarly, SorCS2 deficiency both accelerated the onset and exacerbated the disease motor phenotype in another mouse of model of HD- the zQ175 mice[51]. NMDARs have been shown to play an important role in the regulation of striatum-mediated motor functions. Moreover, both reduced function and genetic variations in the NR2 receptor subunits have been observed in HD patients, with the latter being also linked to the age of onset of motor deficits[51]. In agreement with this, Ma *et al*. (2017)[51] reported decreased cell surface expression of the NMDAR subunit NR2A in the medium spiny neurons (MSNs) in the striatum of zQ175 mice. Subsequent experiments in WT and *Sorcs2*$^{-/-}$ mice identified SorCS2 as NR2A binding partner, which together with VPS35 (a component of the retromer complex) mediates the receptor subunit trafficking to dendrites of the MSNs in the striatum[51]. Additionally, reduced SorCS2 levels were observed in zQ175 mice, as well as in another HD mouse line- R6/2. The decrease was restricted to the striatum and became more pronounced with age. Additionally, SorCS2 was found to interact selectively with the mutant but not the WT form of the

huntington protein prior to the disease onset. As a result, SorCS2 mislocalisation from somato-dendritic compartments to perinuclear regions was observed in both human and mouse HD brain section, potentially interfering with the correct trafficking of the NMDAR subunit NR2A[51]. Meanwhile, sortilin has been functionally linked to Parkinson's disease[52] and FTD[53].

1.1.4.3. The role of the Vps10p-domain receptor family in cognition and neuropsychiatric disorders

Most of the evidence supporting a role of the receptor family in cognition and neuropsychiatric disorders comes from analysis of knockout mice. As described in the sections above, *in vivo* studies have implicated all family members in cellular processes related to synaptic plasticity, and SorLA and SorCS2 have also been shown to be crucial for the development of the dopaminergic system. In addition, recent studies have identified a role of SorCS2 in cellular stress response pathways. All of these processes have been implicated in the aetiology of various neuropsychiatric disorder.

Sorl1$^{-/-}$ mice display impaired object recognition memory from an early age (three months) as they do not distinguish between familiar and new objects. These mice have also shown impaired spatial learning memory in the Morris water maze[43]. Both deficits in object recognition and spatial memory were attributed to the observed increased Aβ plaque deposition[43]. In addition to the memory deficits reported by Capsoni *et al*. (2013)[43], *Sorl1*$^{-/-}$ mice display behavioural traits resembling symptoms of ADHD, such as reduced anxiety levels, hyperactivity and altered response to psychostimulants. It has been suggested that these phenotypes are the result of the decreased nigrostriatal wiring due to the increased GDNF levels observed in these mice as described in section 1.1.4.1[3].

Meanwhile, as described above, *Sorcs2*$^{-/-}$ mice showed abolished NMDA-dependent LTP and LTD as a result of altered TrkB receptor trafficking and BDNF signalling[5]. Behaviourally, these deficits translated into long-term, but not short-term, memory impairments after a weak electrical shock in an inhibitory avoidance test. Barnes test revealed impaired spatial memory in these mice[5]. Additionally, *Sorcs2*$^{-/-}$ mice exhibit ADHD-like behaviours, including hyperactivity, reduced anxiety and increased risk taking in an open field test, as well as dampened response to amphetamines[5,6]. Lack of SorCS2 was also associated with phenotypes reminiscent of affective disorders and schizophrenia (e.g. depression-like response to both tail suspension (TST) and

forced swim tests (FST), and deficits in sensorimotor gating)[5]. According to Glerup *et al.* (2014, 2016)[5,6] these behavioural traits could be explained by dopaminergic hyperinnervation of the frontal cortex and the impaired hippocampal plasticity in response to pro- and mature BDNF found in the *Sorcs2*$^{-/-}$ mice.

More recently, SORCS2 has been linked to cellular pathways involved in neuronal protection against oxidative stress. Increased *SORCS2* expression levels were detected in the CA2 region of the hippocampus of human epileptic brains, as we as in a mouse model of chronic epilepsy- the PTZ (pentylenetetrazol)-kindled mice. Moreover, this increase was particularly apparent in healthy neurons within this area[22]. Meanwhile, when subjected to the PTZ kindling paradigm, *Sorcs2*$^{-/-}$ mice displayed an overall comparable phenotype to the one observed in WT, except for a significant increase in the mortality rate during the course of the experiment. The latter was accompanied by a significant increase in the number of neurons positive for the oxidative stress marker 8-hydroxy-2'-deoxyguanosine (8-OHdG) in the hippocampus of *Sorcs2*$^{-/-}$ mice. This observation, combined with the increased response to oxidative stress detected in primary hippocampal neurons from these mice, suggested that SORCS2 plays a role in the cell defence against oxidative stress[22]. Subsequent experiments showed that SorCS2 is required for the synthesis of the reactive oxygen species (ROS) scavenger- glutathione, through its interaction with the glutamine and cysteine transporter EEAT3 (excitatory amino acid transporter 3). In primary neurons, loss of SorCS2 led to the missorting of EEAT3 from the cell surface and early endosomes to lysosomes, ultimately leading to reduced levels of the transporter and thus altered response to oxidative stress[22]. In addition to its role in oxidative stress response, SORCS2 has been linked to the cell response to alcohol and stress hormones[23,54]. SorCS2 deficiency was also linked to reduced alcohol seeking and withdrawal symptoms *in vivo*. *Sorcs2*$^{-/-}$ mice showed abolished alcohol-induced place preference, which was not the result of impaired contextual memory, and was accompanied by reduced alcohol consumption and lack of alcohol withdrawal symptoms[54]. In addition, *in vitro* studies in SH-SY5Y cells showed increased *SORCS2* expression upon alcohol exposure and withdrawal, as well as following glucocorticoid application[23].

The other members of the receptor family have also been functionally implicated in both cognition and neuropsychiatric disorders. For example, *Sort1*$^{-/-}$ mice displayed memory deficits similar to the ones reported for *Sorl1*$^{-/-}$ mice, but at a later

developmental stage[55]. Knocking out *SORT1* was also associated with reduced immobility time in FST and TST, thought to be a depression resistant phenotype. Further investigation linked the observed phenotype to reduced TREK-1 cell surface levels and increased BDNF content and TrkB expression. In addition to the anti-depressant phenotype, *Sort1*$^{-/-}$ mice also displayed anxiety-like behaviour as it has been assessed in an open plus maze and marble burying test[56]. Meanwhile, unlike *Sorcs2*$^{-/-}$ mice, *Sorcs3*$^{-/-}$ mice showed normal hippocampal LTP but impaired NMDA- and mGluR-dependent LTD, as well as intact acquisition memory, but accelerated extinction of short-term memory in an inhibitory avoidance test. SorCS3-deficient animals also showed deficits in spatial learning and memory, as well as reduced basal synaptic transmission and short-term synaptic plasticity in the hippocampus, potentially as a result of altered AMPA receptor turnover at the post-synaptic density[57]. Lack of SorCS3 was also associated with rapid fear memory extinction in an inhibitor avoidance test, implicating this receptor in aversive memory formation and retrieval[58].

1.1.5. Genetic and expression studies provide further evidence for the role of the Vps10p-domain receptor family in cognition and disease

As discussed in the previous section, all members of the Vps10p-domain receptor family act as binding partners for various pro- and mature neurotrophic factors and synaptic proteins, and knockout mice for these genes recapitulate phenotypes reminiscent of various brain disorders. Thus, in keeping with these functional studies, genetic and expression analyses have provided further evidence for the involvement of the gene family in cognition, learning and memory, as well as neurodegenerative and psychiatric conditions.

1.1.5.1. Genetic and expression studies implicate the Vps10p-domain receptor family in cognition and neurodegenerative disorders, AD in particular

A recent genome-wide association study (GWAS), performed on 1.1 million individuals from a European ancestry, linked *SORL1* (and *SORCS3*) to cognition[59]. It identified SNPs (single nucleotide polymorphisms) within these genes as genome-

wide significant hits for cognitive educational attainment, mathematical ability and highest math class every taken[59]- traits previously shown to correlate with general cognitive performance and to a lesser extend with learning abilities[60,61]. Meanwhile, *SORSC2* shows only suggestive association with cognition at the genome- wide level[62]. The genome-wide significant ($p \leq 5x10^{-8}$) and suggestive ($5x10^{-8} < p \leq 1x10^{-5}$) associations between the sortilins, cognition and cognitive processes are summarised in Table 1.1.

In terms of association of the gene family with neurodegenerative conditions, *SORL1* constitutes the only member that has been linked to AD at a genome-wide level. *SORL1* was first identified as an AD risk gene in a study conducted by Rogaeva *et al.* (2007)[63], who tested several SNPs within the VPS family, including VPS35 and VPS26A, for association with AD. Six SNPs clustered within two haplotype blocks at the 5' and the 3' end of *SORL1* showed nominal association with AD in both the discovery and replication data set ($p<5x10^{-8}$). Moreover, the SNPs located at the 5' block possessed both risk and protective haplotype. The variants identified were non-coding and thus likely to exert their effects via altering *SORL1* transcription and translation[63]. Following from this observation, a more recent study linked the presence of a protective haplotype within the 5' block with increased *SORL1* expression and consequently reduced Aβ production in iPSC-derived neurons in response to BDNF treatment[64].

Subsequent meta-analyses on larger datasets confirmed the association between the SNPs within the two haplotype blocks and AD risk, but also identified additional variants within distinct linkage disequilibrium blocks, conferring the risk of developing AD[65,66]. Interestingly, an additional SNP- rs689021, located within the 5' haplotype block, was shown to be protective against AD within the Caucasian population[65]. Age, sex, ethnic origin and APOE status had no effect on the identified associations[63,65–67]. Additional GWASs further confirmed the association of *SORL1* variants with late-onset AD (LOAD)[68,69,70]. The most recent genome-wide meta-analysis study, combining individuals from three independent cohorts, identified the *SORL1* locus as genome-wide significantly associated with LOAD status ($p=5.57x10^{-11}$), but not with family history of AD ($p=2.81x10^{-6}$). Conducting a meta-analysis on the LOAD and AD-by-proxy (family history of AD) GWASs, where a total of 383,378 controls were compared to 71,880 cases, confirmed the association of the *SORL1* locus with LOAD[70]. In addition, Marioni *et al.* (2018)[71] reported a genome-wide significant

association between variants within *SORL1* and familial history of AD. Polymorphisms in *SORL1* have also been associated with progression from mild cognitive impairment (MCI) to AD[72].

Rare *SORL1* coding variants, including missense, splice site and nonsense mutations, have also been linked to both familial and sporadic AD, as well as early and late onset forms of the disease[41,73–82]. In 2012, Pottier *et al.*[73] reported two missense and five nonsense mutations in *SORL1*, present in autosomal dominant early-onset AD (EOAD) patients, but not in ethnically matched controls. Further investigation revealed allele-specific reduction in *SORL1* expression in peripheral blood cells from a patient carrying the p.Cys1478* mutation. *In silico* analysis of the identified missense mutations showed that they were likely to exert damaging effect, being located within domains known to bind APP and Aβ (e.g. Asn1358Ser mutation within the CR domain and p.Gly511Arg within the Vps10p domain)[73]. In line with these predictions, an independent study in which a mutant version of *SORL1* carrying the p.Gly511Arg variant was overexpressed in SH-SY5Y cells showed reduced Aβ binding to the receptor Vps10p domain, followed by decreased lysosomal degradation[41]. Subsequently, a whole exome sequencing (WES) study, performed by Verheijen *et al.* (2016)[77] on larger cohort of 1255 EOAD cases and 1938 age- and origin- matched controls, identified a significant enrichment of rare missense variants within *SORL1* in patients versus controls. 48% of the rare variants identified were unique to the patient group, and amongst these 80% were predicted to be pathological. Moreover, eight of them- six frameshift variants (p.Thr659Serfs*30, p.Cys752Serfs*21, p.Tyr350fs*, p.Gly447Argfs*22, p.Cys1103Valfs*4, p.Val1747fs*) and two nonsense mutations (p.Arg416* and p.Arg1442*), introduced a premature stop codon. Introducing a premature stop codon is predicted to lead to haploinsufficiency due to non-sense mediated decay (NMD). This was confirmed for one mutation- p.Gly447Argfs*22, which was associated with decrease *SORL1* expression levels in lymphoblast cells, derived from the patient carrier. Treating these cells with cycloheximide increased *SORL1* mRNA levels, suggesting that the observed decrease was due to NMD. None of the identified low-frequency to common variants showed a statistical association with EOAD in this sample set[77]. Similarly, another WES study identified enrichment of rare damaging missense *SORL1* mutations in EOAD patients, especially in those with family history of the disease[80]. Following up on the results from the study conducted by Verheijen *et al.* (2016)[77], Holstege *et al.* (2017)[82] added 640 AD cases (including 320 EOAD case) and 1268

controls from the Netherlands, and classified the rare variants identified based on the predicted severity of effect. The association between a given variant and the risk of AD correlated with frequency, as extremely rare variants increased the AD risk by 12-fold in this sample set. Meanwhile, common variants were not associated with AD, even if they were predicted to be highly damaging[82].

In addition, Vardarajan *et al.* (2015)[75] reported the presence of rare, coding mutations in *SORL1*, which were predicted to be damaging and segregating with LOAD. The effect of three of the identified mutations was functionally assessed in HEK293 cells by overexpressing mutated version of the *SORL1* gene. Introducing the p.Gly270Lys and Thr947Met mutations led to an increased secretion of $A\beta_{40}$ and $A\beta_{42,}$ while overexpressing the Ala528Thr variant increased $A\beta_{42}$, but not Aβ40, levels. All three mutations were associated with elevated sAPPα and sAPPβ production, as well as weakened SORLA binding to APP. However, only the Thr947Met variant influenced SORLA cell surface levels, suggesting that the other two mutants might affect APP processing by disrupting its interaction with SORLA at the cell surface[75]. Interestingly, all three variants were located within close proximity to SNPs initially identified by Rogaeva *et al.* (2007)[63]. Similarly, Cuccaro *et al.* (2016)[76] and Gomez-Tortoza *et al.* (2018)[81] found additional rare variants segregating with both LOAD and EOAD. In keeping with the findings published by Vardarajan *et al.* (2015)[75], functional analysis of the novel *SORL1* variants identified by Cuccaro *et al.*[76] (p.Thr588Ile and pThr2134Met) confirmed their impact on APP trafficking and Aβ production.

In summary, the studies described above have identified rare coding variants, scattered throughout the different domains of SORLA and predicted or functionally validated to have negative consequence in terms of SORLA's role in APP trafficking and processing into Aβ. However, these studies applied diverse strategies in terms of patient inclusion criteria, variant identification and assessment. In order to overcome these limitations, a meta-analysis study published this year combined and reanalysed the data from five non-overlapping WES studies[74]. The variants were filtered according to their allele frequency and predicted damaging effect. Protein truncating variants conferred the highest AD risk, and missense mutation classified as damaging by three independent tools (SIFT, Mutation Taster and Polyphen2) were associated with AD at an exome-wide significant level, independent of their allele frequency. These associations were stronger when assessed only for EOAD cases (as opposed to a combined sample of EOAD and LOAD). Variants located within the Vps10p

domain, CR domain or fibronectin type III domain showed stronger association with AD risk compared to mutations located in other domains[74].

SORCS2 does not show association with AD, but shows suggestive association ($5x10^{-8} < p \leq 1x10^{-5}$) with hippocampal sclerosis, an AD neuropathological feature[83]. With regard to the rest of the family members, *SORCS3* has shown suggestive association with LOAD risk in a family-based GWAS meta-analysis[84], and a non-coding *SORT1* variant has been linked to decreased risk of developing AD[85]. In addition to AD, rare variants in *SORT1* and *SORCS3* have been associated with frontotemporal dementia (FTD)[86] and multiple sclerosis[87], respectively. Genome-wide significant ($p \leq 5x10^{-8}$) and suggestive ($5x10^{-8} < p \leq 1x10^{-5}$) associations linking members of the Sortilin family to neurodegeneration are summarised in Table 1.1.

In terms of expression, reduced SORLA protein levels have been detected in post-mortem brain tissue derived from AD patients[39,88]. Reduced *SORL1* mRNA levels have also been found in brain areas of cognitively intact individuals that show early signs of AD-related neuropathology[89]. Additionally, two variants in LD, located within *SORL1* 3' end, were significantly associated with reduced mRNA and protein levels in AD patient brain speciments[90]. *SORL1* risk variants have also been shown to reduce *SORL1* expression during childhood and early adolescence, as well as to predict endophenotypes of AD (e.g. decreased white matter fractional anisotropy and increased amyloid pathology in post-mortem brain) in individuals who were dementia-free at the time of the study[91]. Reduced *SORCS1* expression was also detected in the amygdala, but not the cerebellum (a brain region spared in AD) of AD patients[47].

Table 1.1. GWAS data linking the sortilins to cognition and neurodegeneration. The columns show a given family member, the trait with which it has been associated, the identified SNPs and their significance at the genome-wide level. The far-right column provides a reference to the respective study.

Gene	Trait	SNP	p-value	PUBMED ID
SORL1	Alzheimer's disease in APOE e4-carriers	rs11218343	$9x10^{-6}$ (suggestive)	25778476
	Alzheimer's disease (family history)	rs11218343	$1x10^{-11}$	30617256
		rs11218343	$3x10^{-6}$ (suggestive)	30617256
		rs11218343	$2x10^{-6}$ (suggestive)	29777097
	Alzheimer's disease (late onset)	rs11218343	$1x10^{-14}$	24162737
		rs11218343	$6x10^{-11}$	30617256
	Educational attainment (self-reported)	rs12146618	$3x10^{-9}$	30038396
	Mathematical ability (highest math class taken)	rs12146618	$4x10^{-11}$	30038396
		rs11218447	$1x10^{-8}$ (suggestive)	
	Mathematical ability (self-reported)	rs1448133	$4x10^{-9}$	30038396
	Non-word reading (dyslexia)	rs556349	$2x10^{-6}$ (suggestive)	30741946
	Word reading (dyslexia)	rs556349	$6x10^{-6}$ (suggestive)	30741946
SORT1	Cognitive ability	rs11102974	$1x10^{-9}$	29186694
SORCS1	-	-	-	-
SORCS2	General cognitive ability	rs192716625	$7x10^{-6}$ (suggestive)	29844566
	Hippocampal sclerosis	rs62277617	$8x10^{-6}$ (suggestive)	25188341

Gene	Trait	SNP	p-value	PUBMED ID
SORCS3	Alzheimer's disease and age of onset	rs117792039	$2x10^{-7}$ (suggestive)	26830138
	Cognitive ability	rs1947988	$1x10^{-8}$	29186694
		rs1947988	$2x10^{-8}$	29326435
		rs7895991	$3x10^{-8}$	29942086
		rs1947988	$8x10^{-7}$ (suggestive)	29186694
		rs1484246	$1x10^{-6}$ (suggestive)	29844566
		rs902305	$9x10^{-6}$ (suggestive)	29844566
	Cognitive performance	rs3896224	$5x10^{-14}$	30038396
		rs790647	$3x10^{-12}$	
		rs55680960	$3x10^{-10}$	
	Educational attainment (self-reported)	rs902305	$2x10^{-22}$	30038396
		rs790647	$4x10^{-18}$	
		rs12252384	$5x10^{-14}$	
		rs11599236	$1x10^{-13}$	30038396
		rs7914674	$2x10^{-10}$	
		rs10509790	$3x10^{-10}$	
		rs11192147	$3x10^{-9}$	
		rs1484246	$4x10^{-9}$	30595370
		rs1869165	$1x10^{-13}$	

Gene	Trait	SNP	p-value	PUBMED ID
SORCS3	Mathematical ability (highest math class taken)	rs10786832	$6x10^{-20}$	30038396
		rs17118088	$5x10^{-18}$	
		rs902305	$7x10^{-12}$	
		rs11192147	$5x10^{-11}$	
		rs12356045	$3x10^{-10}$	
		rs703482	$4x10^{-9}$	
		rs11192147	$7x10^{-9}$	
	Mathematical ability (self-reported)	rs2864034	$2x10^{-16}$	30038396
		rs2864034	$1x10^{-12}$	
		rs10400054	$2x10^{-11}$	
		rs4537697	$2x10^{-9}$	
		rs12356045	$2x10^{-8}$	

1.1.5.2. Genetic studies implicate the Vps10p-domain receptor family in psychiatric disorders

SORL1, *SORCS2* and *SORCS3* have been genetically associated with risk taking behaviours. A GWAS on over one million individuals from UK Biobank, identified a SNP within *SORCS2* as a genome-wide significant hit for general risk-taking behaviour, and SNPs in *SORL1*- for alcohol consumption and smoking status[92]. Interestingly, these risky behaviours correlated genetically with traits like neuroticism, extraversion and openness to new experience, as well as with psychiatric phenotypes such as ADHD, bipolar disorder and schizophrenia[92]. Additionally, a recent GWAS identified *SORCS2* as a genome-wide significant hit for alcohol withdrawal severity. Moreover, the risk haplotype was mapped to a regulatory region, known to be specifically active in the hippocampus and modulated by stress hormones[23]. Similarly, two independent studies identified *SORCS3* as a GWAS hit for both smoking status[93] and initiation[94].

SNPs within *SORCS2* have been also associated at a suggestive level ($5x10^{-8} < p \leq 1x10^{-7}$) with various psychiatric conditions and phenotypes, including anorexia nervosa[95], attention deficit in ADHD[96], long-term treatment outcome in bipolar disorder (BD)[97] and neuroticism[98]. Additionally, one GWAS on pooled DNA from two independent case-control sample sets[99] and two candidate gene studies[100,101] further linked *SORCS2* to the aetiology of BD. However, amongst all members of the receptor family, *SORCS3* shows the strongest association with psychiatric illness, in particular depression and associated phenotypes[102–105]. The GWAS-based links between members of the sortilin family and psychiatric illness are summarised in Table 1.2.

Table 1.2. GWAS data linking the sortilins to psychiatric disorders. The columns show a given family member, the trait with which it has been associated, the identified SNPs and their significance at the genome-wide level. The far-right column provides a reference to the respective study.

Gene	Trait	SNP	p-value	PUBMED ID
SORL1	Alcohol consumption	rs530916	$1x10^{-12}$	30643251
		rs485425	$5x10^{-12}$	30643258
	Mood disorder in prion disease	rs11218343	$9x10^{-6}$ (suggestive)	25897833
	Smoking initiation	rs568599	$5x10^{-17}$	30643251
		rs540860	$6x10^{-12}$	
		rs1944689	$1x10^{-8}$	
	Smoking status	rs568599	$1x10^{-10}$	30595370
		rs568599	$3x10^{-10}$	30643258
		rs1944688	$4x10^{-8}$	30643258
SORT1	-	-	-	-
SORCS1	Methadone dose in opioid dependence	rs6584756	$4x10^{-6}$ (suggestive)	28115739
SORCS2	Anorexia nervosa	rs13125782	$9x10^{-7}$ (suggestive)	28494655
	Attention function in attention deficit hyperactive disorder	rs4689642	$4x10^{-7}$ (suggestive)	26174813

Gene	Trait	SNP	p-value	PUBMED ID
SORCS2	Depressive and manic episodes in bipolar disorder	rs16840900	$4x10^{-6}$ (suggestive)	26297903
	Neuroticism	rs34369999	$6x10^{-6}$ (suggestive)	27089181
	Risk taking behaviour	rs3846425	$1x10^{-8}$	30643258
	Response to antidepressants (symptom improvement)	rs17382228	$2x10^{-6}$ (suggestive)	29160301
SORCS3	Attention deficit hyperactivity disorder	rs11591402	$3x10^{-9}$	30478444
	Autism and major depressive disorder	rs61867293	$1x10^{-6}$ (suggestive)	30804558
	Depression	rs1021363	$4x10^{-23}$	30718901
		rs11599236	$3x10^{-11}$	29942085
		rs61867293	$7x10^{-10}$	29700475
		rs1961639	$4x10^{-9}$	29942085
		rs10786831	$8x10^{-9}$	27479909
		rs1021363	$1x10^{-8}$	29662059
		undefined	$2x10^{-8}$	29942085
		rs11599236	$8x10^{-8}$ (suggestive)	29662059
		rs7074335	$2x10^{-6}$ (suggestive)	27089181

1.2. DNA DSB formation and repair in neuropsychiatric disorders

1.2.1. Maintenance of genome integrity in the nervous system

As it is isolated from the environment, the central nervous system (CNS) is protected from the most common exogenous sources of DNA damage, such as. UV radiation and tobacco smoke. However, endogenous DNA damage occurs both during development and in the mature nervous system. During early development, and more specifically neurogenesis, the high replication rate of neuronal precursor cells results in replication-associated DNA damage. Meanwhile, in mature neurons, the extremely high energy demands, and metabolic activity associated with high transcriptional rates lead to the production of ROS. ROS, together with their reaction products, such as reactive nitrogen species and lipid peroxidation products, constitute the main source of DNA damage in the brain[106]. Recently, neuronal activity has been also linked to DNA damage[107]. The maintenance of cellular genome integrity is vital for long-lived cells, such as neurons with no or limited regeneration from precursor cells. Accumulation of excessive and/or unrepaired damage can lead to neuronal loss.

In the nervous system, there are four main DNA damage response (DDR) pathways, each of which responds to a particular type of damage[106] (Figure 1.4.). The base excision repair (BER) pathway detects and corrects base modification caused by the production of ROS. Meanwhile, the nucleotide excision repair (NER) mechanism is responsible for the repair of DNA lesions resulting in from the exposure to UV radiation or chemicals. Single-stranded DNA breaks (SSBs) can arise from damage of the DNA backbone caused by ROS, as well as intermediates of the BER pathway, and are repaired by the single strand break repair (SSBR) pathway. DNA double-stranded breaks (DSBs), the type of DNA lesions discussed in this thesis, can be repaired by either non-homologous end joining (NHEJ) or homologous recombination (HR). DSBs can be generated in various ways ranging from unrepaired SSBs and replication stress (i.e. exposure to ionizing radiation or chemotherapeutic agents)[108]. Unresolved, they are amongst the most dangerous DNA lesions, as they can cause large chromosome rearrangements leading to cell death or tumorigenesis. Thus, their repair is crucial for cell survival. The choice of the repair pathway depends on the presence of specific proteins and the cell cycle. For example, HR occurs only during the S- or G2-phase of the cells cycle as it requires the presence of a sister chromatid

serving as a template for error-free repair. Meanwhile, NHEJ, an error-prone process, involves processing and direct ligation of the damaged DNA ends, without the need for a repair template. Thus, NHEJ can take place at any stage of the cell cycle and is the only pathway responsible for the repair of DSBs in post-mitotic neurons.

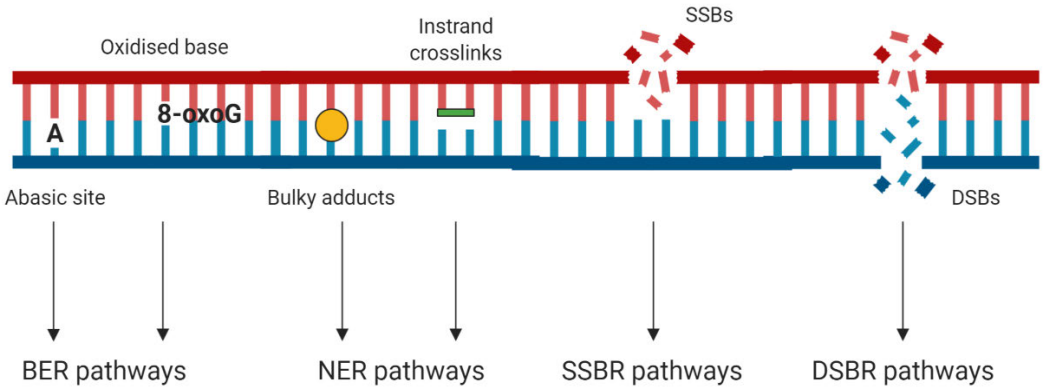

Figure 1.4. Types of DNA damage and the corresponding DNA repair pathways. Single base damage (e.g. abasic site or oxidised based) are generated by ROS, nitrogen species and other metabolic products. This type of DNA lesions is repaired by the BER (base excision repair) pathway. Modifications that destroy the double helix (e.g. bulky adducts and instrand crosslinks) are repaired by the NER (nucleotide excision repair) pathway. These lesions occur as a result of UV radiation, exposure to chemicals and ROS. SSBs (single-stranded DNA breaks) can be caused by ROS activity, ionizing radiation and chemotherapeutic reagents and are repaired by the SSBR (single-strand break repair) pathways. Meanwhile, DSBs (double-stranded DNA breaks), caused by ionizing radiation, chemotherapeutic reagents, altered TopoII activity and neuronal activity, are repaired via the DSBR (double-strand break repair) pathways. [Created with BioRender]

1.2.2. γH2A.X as a DSB marker in the CNS

The hallmark of DSB formation is the phosphorylation of a major core histone H2A variant, H2A.X, on Ser139 (termed γH2A.X). Its role in DNA damage and DSB repair was first demonstrated by Rogakou *et al.* (1998)[109], who showed that exposure to ionizing radiation led to rapid (as early as 10 min post-irradiation) phosphorylation of H2A.X both *in vivo* and *in vitro*. Subsequently, it was shown that H2A.X phosphorylation appeared as foci, and that their number corresponded directly to the number of DSBs[110]. The rapid induction of γH2A.X foci formation, together with the

1:1 ratio between the number of γH2A.X foci and DSBs has made antibodies against γH2A.X the 'gold standard' for detecting DSBs[111]. Subsequent work has shown that γH2A.X serves as a docking site for various DNA damage response factors, including the Mre11 complex, 53BP1 (p53-binding protein 1) and BRCA1[112]. Alongside providing a platform for the DDR machinery, there is evidence that γH2A.X also facilitates DNA repair by binding to the broken DNA ends and thus keeping them in close proximity[113]. Apart from radiation, phosphorylation of H2A.X can be induced by other genotoxic stressors, such as UV, hydroxyurea, topoisomerase inhibitors, environmental and chemical agents[112].

In the mouse brain, phosphorylation of H2A.X was detected throughout the entire rosto-caudal axis and in all stages of mouse development- prenatal, postnatal, adult and senescence[111]. In embryos, γH2A.X foci were present in brain areas known for their neurogenic capacity: the subgranular zone (SGZ), rostro migratory stream (RMS), olfactory bulb (OB) and third ventricle. In these areas γH2A.X foci were primarily found in proliferating cells. Meanwhile, in postnatal mice, the strongest γH2A.X signal was seen in the cerebellum, reaching its peak at P5-10. A clear reduction in the number of γH2A.X-positive nuclei was observed in adult and aged mice (two years). After P15, the γH2A.X signal intensity decreased in the cerebellum, while increasing in the cerebral cortex. Here, γH2A.X was primarily found in mature neurons and most likely as a result of the DNA damage response system activation rather than proliferation. In the hippocampus, γH2A.X foci were observed primarily in the SVZ of the dentate gyrus in mice aged between P10 and P60[111].

As reported in other tissues, ionizing radiation of embryonic mouse brain led to the rapid phosphorylation of H2A.X in both neurons and neuronal progenitor cells. Twenty-four hours later, the observed γH2A.X foci were completely abolished in neurons, but not in precursor cells, which displayed apoptotic features[114]. γH2A.X has also been shown to provide an early and sensitive marker of neuronal endangerment upon non-lethal insults[115]. In rat cortical neurons, γH2A.X was detected following non-cytotoxic stimulation of the ionotropic glutamate receptors- NMDA and AMPA receptors[115]. The observed levels of γH2A.X induction were comparable to those detected after exposure to other conventional DNA damaging agents (e.g. ionizing radiation and hydrogen peroxide) and were shown to be at least partially dependent on ROS production. The majority of the detected γH2A.X foci (>50%) co-localised with the DSB repair protein Mre11, suggesting that at least a proportion were DNA

damage-dependent, as well as showing that the DNA repair pathway was activated. The authors of the paper suggested that the proportion of foci that failed to co-localise might have represented either transient co-localisation due to the rapid nature of the DNA damage detection and repair process or a DSB-independent function of γH2A.X (e.g. regulation of gene expression by modification the chromatin structure)[115].

A similar pattern of γH2A.X formation was also observed *in vivo*. In adult rats, recurrent individual seizures following injection with kainic acid (KA) led to a rapid (within 30 min) increase in the number of γH2A.X positive foci in multiple brain regions, including the hippocampus and the entorhinal cortex. Allowing the animals to enter status epilepticus resulted in higher γH2A.X levels after only five minutes exposure, with the highest levels reached after 120 minutes. This was the case in all brain areas examined, except for the dentate gyrus, which appeared to be resistant to seizure-induced DNA damage. Interestingly, the duration of both recurrent seizures and status epilepticus, required for γH2A.X induction, was insufficient to cause neuronal death[116].

1.2.3. DNA DSB formation- physiological function in the CNS and in neurodegenerative disorders

In the developing brain, DNA DSB formation has been associated with DNA synthesis and proliferation of neural stem cells as shown by the co-localisation of γH2A.X with BrdU and a phosphorylated form of histone H3 (pHH3), respectively [111]. In the adult brain, γH2A.X foci formation has been shown to be involved in the regulation of adult neurogenesis[117]. More recently, DNA DSB formation has been linked to physiological brain activity[107] and activity-induced transcription of early- response genes[118]. Suberbielle *et al.* (2013)[107] linked a natural behaviour, exploration of a novel environment, to increased DSB formation. Levels of DSB formation were measured as the number of neurons in various brain regions, linked to learning and memory including the dentate gyrus of WT mice, that were positive for γH2A.X[107]. Follow-up experiments showed that the increase in the number of γH2A.X-positive nuclei was the result of neuronal activity, and not stress-induced release of hormones (as adrenalectomy had no effect on the observed phenotype). Meanwhile, both exposure to visual stimuli and direct optogenetic stimulation led to elevated DSB formation in the visual cortex and the striatum on WT mice[107]. Subsequently, Madabhushi *et al.*

(2015)[118] confirmed these observations both *in vitro* and *in vivo*. Pharmacological and electrophysiological stimulation of primary neurons and hippocampal slices, respectively, led to DSB formation accompanied by increased transcription of *FOS* and *NPAS4*[118]. Further ChIP-sequencing experiments, following neuronal stimulation with NMDA, showed that the γH2A.X signal occurred mostly at the promoters of early-response genes, including *Fos* and *Npas4*[118]. Similarly to the treatment with NMDA, treatment of primary neurons with etoposide, but not other DSB-inducing agents (i.e. radiomimetic drugs and DDR inhibitors), resulted in increased expression of the same early-response genes. All together these results suggested that the increased expression of early-response genes depended on the formation of DSBs as a results of Topoisomerase II (TopoII) activity. ChiP-seq, knockdown and immunoprecipitation experiments confirmed the role of TopoIIβ in activity-induced DSB formation in neurons. Moreover, utilising the CRISPR/Cas9 genome editing technique, Madabhushi *et al.* showed that it was precisely the formation of DSBs that triggered the expression of the early-response genes[118].

In addition to the increased DSB formation in WT mice following exploration of a novel environment, Suberbielle *et al.* (2013)[107] reported elevated baseline DSB levels in hAPP mice- a mouse model of AD, carrying a human APP (hAPP) transgene[107]. When subjected to the same exploratory task, these mice showed an increase in the number of neurons positive for γH2A.X comparable to the WT mice, but their capacity of repairing these DNA breaks was reduced. hAPP mice show aberrant network activity, and interventions that counteract this dysfunction (e.g. genetic ablation of tau or administration of anti-epileptic drugs) have been associated with improved synaptic and cognitive functions in these mice. Similarly, crossing hAPP mice with a line knockout for *MAPT* (which encodes tau) or treating them with levetiracetam (an anti-epileptic drug) reduced the number of γH2A.X-positive neurons both at baseline and following recovery from the novel environment[107]. Incubating WT primary neurons with Aβ oligomers led to an increase both in the number of cells positive for γH2A.X and γH2A.X levels. In agreement with the observations from the hAPP mice, treating these cells with a sodium channel blocker (TTX) or with AMPAR or NMDAR antagonists abolished the Aβ-induced increase in DSB formation. Based on these results, Suberbielle *et al.* (2013)[107] hypothesised that the effect of Aβ on DSB formation depended on neurotransmitter release as a results of action potential (AP) generation, as well as on the activation of NMDA receptors following AMPAR-mediated depolarisation of the synaptic membrane. Further experiments, involving culturing

neurons under conditions stimulating either NR2A- or NR2B-containing NMDARs, showed that the former is required for DSB repair, while the activity of the latter is necessary for DSB formation both in the presence and absence of Aβ oligos. While activation of NR2B-containing NMDARs is associated with increased production of ROS, treatment with the ROS-scavenger Euk-134 had no impact on γH2A.X foci formation under these conditions[107].

A follow-up study attributed the accumulation of DSBs in the neurons of hAPP mice to defective DNA repair machinery[119]. Reduced levels of the DSB repair factor BRCA1 were detected in hAPP mice and were also seen in post-mortem brain tissue from AD patients. Knocking down *BRCA1* led to increased formation and compromised repair of neuronal DSBs both in WT and hAPP transgenic mice. This was accompanied by learning and memory deficits, which were more pronounced in the hAPP mice. In WT mice, knock-down of *BRCA1* was also associated with changes in neuronal morphology (e.g. smaller cell bodies and reduced dendritic branching) and function (e.g. altered AP frequency and LTP deficits). Meanwhile, increased neuronal activity following an exploratory activity (*in vivo*) or NMDA receptor stimulation (*in vitro*) induced *BRCA1* expression. *In vivo*, this increase was less pronounced in the hAPP mice, potentially accounting for the compromised repair of activity-induced DSBs observed in them[119]. Additional post-mortem brain studies provide further evidence for the potential role of deficient DNA repair mechanisms in the accumulation of DNA damage observed in AD patients [120,121].

1.2.4. DNA DSB formation in neuropsychiatric disorders, with focus on AD

In contrast to the temporary acquisition of DNA DSB formation following neuronal activity, accumulation of this type of DNA lesion has been implicated in aging, neurodegenerative and neurological conditions, including AD, ataxia-telangiectasia, ataxia-telangiectasia-like disorder and Nijmegen breakage syndrome[122,123]. With respect to AD, initial post-mortem brain studies detected increased levels of DNA breaks (single- and double- stranded) in the hippocampus and the cerebral cortex of AD patients[124,125]. A more recent study reported increased expression of γH2A.X, parallel to that of a protein kinase (DNA-PK) involved in NHEJ, in both neurons and astrocytes of aging individuals[126]. These markers of DNA damage did not, however,

correlate with Aβ and tau pathology or Braak stage[126]. In contrast, further studies identified increased levels of DNA damage amongst a subset of early Braak stage individuals, who showed lower Mini-mental state examination (MMSE) scores than the rest[127]. Elevated γH2A.X and DNA-PK levels occurred independently of ApoE status, vascular pathology or accumulation of Aβ plaques, and were associated with increased activity of the senescence-associated protein β-galactosidase, but not with the further activation of senescence pathways[127]. These results suggested that the elevated levels of DDR contributed to neuronal dysfunction and cognitive decline independently of the development of AD pathology. A more recent paper provided further evidence supporting the notion that accumulation of DSBs in susceptible neurons and glia cells commences during early AD stages and persists throughout the disease progression. Using immunohistochemistry, Shanbhag *et al.* (2019)[128] detected an increased proportion of γH2A.X-positive neurons and astrocytes in the frontal cortex and the CA1 region of the hippocampus of both MCI and AD cases compared to cognitively healthy controls.

Elevated levels of DNA damage have been also reported for a number of psychiatric disorders, including MDD, bipolar disorder and schizophrenia[129]. It is hypothesised that this is primarily the result of increased oxidative stress, as the DNA lesions detected were positive for 8-OHdG, which is one of the most widely used biomarkers of DNA damage due to oxidative stress[129]. Similarly, increased oxidative stress has been reported in post-mortem brains from MCI and AD patients, and elevated levels of 8-OHdG have been detected in brains from individuals with late-stage AD[130]. Usually, DNA lesions caused by oxidative stress are repaired by the BER or the NER pathways, as discussed in section 1.2.1. However, unrepaired these lesions can accumulate, forming oxidative clustered DNA lesions (OCDLs) and eventually leading to the formation of DSBs[131]. There is conflicting evidence of whether γH2A.X provides a reliable measure for oxidative stress-induced DNA damage. *In vitro*, treatment with hydrogen peroxide, a commonly used oxidative stress inducing reagent, led to an increase in the levels of γH2A.X staining[132]. However, this signal was distributed throughout the nucleus rather than appearing at distinct foci, as observed following ionizing radiation. Moreover, the detected γH2A.X staining did not always co-localise with 53BP1 or ATM (the main kinase that phosphorylates H2A.X in response to DSB formation) and knocking out *XRCC4* (a major component of the NHEJ pathway) had no effect on γH2A.X elimination. Taken together, these results suggested that the observed induction in γH2A.X, as a result of increased oxidative stress, was not the

result of DSB formation[132]. In contrast, another study found that knocking out major components of the NHEJ and HR *in vitro* conferred susceptibility to low doses of hydrogen peroxide[131]. This was accompanied by an increase in the formation of 53BP1 foci, partially overlapping with γH2A.X signal, providing evidence that even low levels of oxidative stress can trigger the formation of DSBs[131].

1.3. LUHMES as a model system

1.3.1. The LUHMES neuronal cell line

The LUHMES (Lund Human Mesencephalic) cell line is a karyotypically normal human foetal mesencephalic cell line conditionally immortalised with the v-myc oncogene[133]. Proliferation of these neuronal precursor cells can be terminated by adding tetracyclin, thus halting v-myc expression. Subsequent addition of neurotrophic growth factors (e.g. GDNF) results in differentiation into post-mitotic dopaminergic neurons within five days. Five to six days after initiating differentiation, the expression of stem cell and neuronal precursor markers, such as SOX2 and PAX3, is significantly reduced, while that of neuronal markers, such as synaptophysin, synapsin and PSD-95, is strongly upregulated. While the expression of most neuronal markers peaks at day 6 and remains stable thereafter, the electrophysiological properties of differentiated LUHMES cells have been shown to progressively increase with the majority of cells generating spontaneous APs after 10-12 days of differentiation[133]. In addition to expressing pre- and post-synaptic, and dopaminergic markers, differentiated LUHMES cells have also been shown to express proteins relevant to AD, including APP, Aβ, sAPPβ, tau, p-tau and γ-secretase[134]. Members of the APP gene family, as well as genes involved in APP processing (e.g. α-, β- and γ-secretases) and turnover (e.g. APOE and LRP1) were expressed in proliferating cells, with the expression of some of them increasing dramatically over the course of differentiation. At the protein level, a shift in the presence of immature to mature forms were observed for BACE and APP. Similarly, sAPPβ and Aβ secretion were upregulated upon differentiation, with the latter being triggered by GDNF and activation of its co-receptor RET[134].

1.3.2. Relevance of the dopaminergic system in respect to the biology of SORLA and SORCS2

As described above, LUHMES neurons are dopaminergic, and despite not being the most obvious choice of neuron type for studying processes related to learning and memory and Aβ pathologies, they were an appropriate model system for this study for the reasons described below. First, LUHMES constitute a human neuronal cell line that can be easily modified genetically[135], which provides the advantage of studying human, rather than mouse primary neurons. Second, LUHMES cells can be rapidly differentiated into homogenous population of electrophysiologically active dopaminergic neurons within less than two weeks. This makes them a more cost and time effective *in vitro* system than iPSCs (induced pluripotent stem cells), which can take between 6-15 weeks until differentiated into neurons, depending on the protocol and the level of maturity required[136]. Moreover, both SorCS2 and SorLA have been shown to regulate differentiation and proliferation of dopaminergic neurons via regulation of BDNF and GDNF activity, respectively[3,6].

Additionally, with regards to AD, studies of both transgenic mouse models of AD and patients provide evidence for the involvement of dopaminergic neurons and brain circuits in the disease pathology. For example, in a cohort of healthy individuals and patients with MCI and AD, functional connectivity between the VTA and the hippocampus was correlated with hippocampal volume and memory performance[137]. Moreover, parkinsonism symptoms, such as bradykinesia, face masking, tremors, gait disturbances, are commonly seen in AD patients and are often attributed to dysfunction of the dopaminergic nigrostriatal pathways[138]. The presence of parkinsonian features and Lewy body pathology has been reported in individuals diagnosed with AD and carrying rare mutations in *SORL1*[76]. Additionally, the AD mouse model APP/PS1 was shown to display dopaminergic pathology in the striatum and substantia nigra, which was closely associated with age-dependent amyloid deposition[139]. Recent analysis of another murine model of AD (Tg2576) showed significant, progressive loss of dopaminergic neurons in the VTA, prior to Aβ deposition[140]. Moreover, the accompanying reduction in dopaminergic innervation to the hippocampus was correlated with impairments in synaptic plasticity, memory performance and reward processing in these mice[140].

1.3.3. Using CRISPR/Cas9 genome editing to introduce mutations in the LUHMES neuronal cell line

CRISPR has recently become one of the most popular approaches for genome editing due to its simplicity, high specificity and efficiency compared to other editing technologies, such as zinc-finger nucleases and transcription activator-like effector nucleases[141]. Derived from the bacteria immune system, CRISPR comprises of two components- a 'guide' RNA (gRNA) and a non-specific endonuclease- Cas9, which is guided to the target DNA by the gRNA via Watson-Crick base pairing[141]. The target DNA can be any 20-nucleotide sequence, preceding a species-specific 3bp 5' PAM sequence, which is necessary for target binding. The Cas9 protein and the gRNA form a riboprotein switching the Cas9 protein from an inactive to an active, DNA-binding conformation. The Cas9:gRNA complex then binds to any putative DNA sequence with a PAM site and if the gRNA and the target DNA match, the former will anneal to the later in a 3' to 5' direction. At this stage, the Cas9 protein undergoes a second conformational change which allows the endonuclease to cleave both strands of the target DNA. This results in the formation of a DSB that could be repaired either via NHEJ or HDR (homologous-directed repair). In post-mitotic cells, the NHEJ is the more active, but also error-prone repair mechanism[141]. Activation of this pathway can result in insertions, deletions or frameshift mutations at the DSB site, potentially leading to a gene knockout. While the NHEJ repair mechanism is used for disrupting the open reading frame of a gene, the HDR system can be implemented in generating specific nucleotide changes ranging from a single nucleotide substitution to large insertion. In order to do this, a DNA repair template containing the desired sequence must be delivered alongside the gRNA and the Cas9 protein. The repair template can take the form of a single- or double-stranded oligonucleotide, as well as a double-stranded DNA plasmid[141] (Figure 1.5).

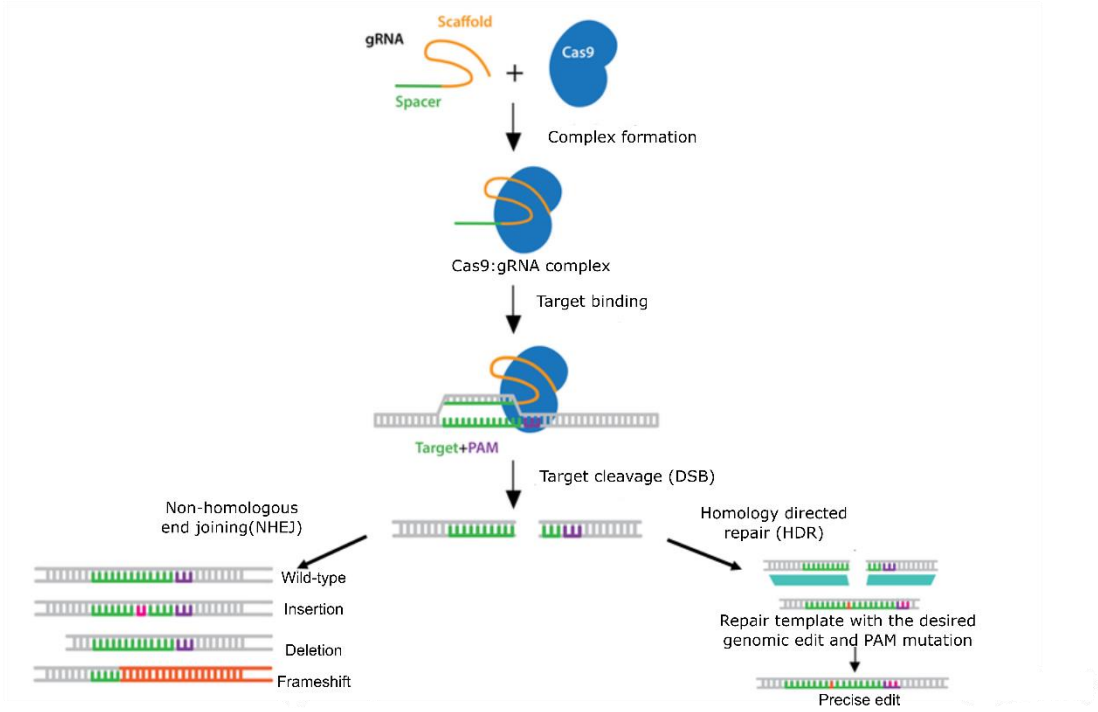

Figure 1.5. Genome editing using CRISPR/Cas9. The first step is the formation of a riboprotein, following the gRNA association with the Cas9 protein. The Cas9:gRNA complex then bind to the target DNA, which results in the formation of a DSB that could be repaired either via NHEJ or HDR. Being error-prone, the NHEJ can lead to insertions, deletions or frameshift mutations at the DSB site, potentially generating a knockout. Meanwhile, the HDR system can be used to generate specific nucleotide changes by adding a DNA repair template containing the desired sequence. [Adapted after https://www.addgene.org/crispr/guide/].

One of the main considerations when designing CRISPR/Cas9-mediated genome editing experiments is ensuring the specificity of the system. This is mostly determined by how unique the gRNA target sequence is compared to the rest of the genome. Usually, a gRNA target sequence will show partial homology with additional sites, known as off- target sites, throughout the genome. The available online tools for gRNA design take this into account and can rank the gRNA target sequences for a given region of interest based on their predicted on- and off-target activity. The specificity of the system can be further increased by using modified versions of the Cas9 endonuclease, such as the Cas9 nickase (Cas9n), which generates single strand nicks in the DNA. Thus, a double nickase is required for introducing a DNA DSBs at the target site, which increases the target specificity dramatically. More recently, high fidelity Cas9 systems have also been developed, in which mutagenesis has been

employed to minimize off-target activity by weakening the enzyme interaction with the DNA or increasing its proofreading and target discrimination activity.

To date the CRISPR-Cas9 genome-editing system has been utilised in creating both *in vivo* and *in vitro* models for studying human disorders, including AD[142]. Moreover, this strategy has already been successfully implemented in LUHMES cells where it has been used to introduce missense mutations in the *MECP2* and *EEF1A2*- genes, which have been associated with Rett syndrome and neurodevelopmental delay with intellectual disability, respectively[135].

1.4. Main project aims

The aims of the project discussed in this thesis stem from a few work streams. First, as discussed above, DNA DSB formation, associated with γH2A.X foci occurrence, has been shown to play a role in the aetiology of many neurological and neuropsychiatric conditions, as well as in processes related to learning and memory. In mice, increase in neuronal activity as a result of an exploratory activity resulted in a transient increase in DSB formation, repaired within 24 hours. In the hAPP mice, this phenotype was exacerbated, and *in vitro* experiments revealed that the Aβ-induced DSB formation was dependent on NR2B-containing NMDAR activativity[107] Second, genetic and functional studies have implicated the sortilins in cognition, as well as in various neurodegenerative and psychiatric disorders. Third, knocking out members of the receptor family have been associated with increased APP processing and Aβ accumulation (SORLA and sortilin *in vivo* and *in vitro*; SORCS1-3 *in vitro*). More recently SorCS2 has also been shown to play a role in NMDAR subunit trafficking[51]. With this in mind, the aims of my PhD project were to:

1) Investigate the levels of DNA DSB formation in WT, *Sorcs2*$^{-/-}$ and *Sorl1*$^{-/-}$ mice, exposed to a novel environment task, previously reported to introduce a transient increase in the number of DSBs in brain areas important for learning and memory.
2) Knock out *SORCS2* and *SORL1* in the human neuronal cell line LUHMES and assess the effect of the introduced mutations on DSB formation.
3) Explore potential mechanisms that could account for the observed effect of knocking out *SORCS2* on DSB formation *in vivo* and *in vitro*.
4) Introduce the rare EOAD- associated, *SORL1* missense mutation G508S and examine its impact on *SORL1* expression and Aβ production.

Chapter 2 Materials and Methods

2.1. Materials

2.1.1. Buffers, solutions and gel loading dyes

Table 2.1. Table listing all buffers, solutions and loading dyes used.

Buffer/ Solution/ loading dye	Component	Volume/ weight/ concentration
Laemmli protein sample buffer (2x)	Tris (1M, pH 6.8) Glycerol SDS (20% solution) DTT (1M) Bromophenol Blue Qs (Quantity sufficient) to 10ml dH_2O	1ml (100mM) 2ml (20%) 2ml (4%) 2ml (200mM) 0.2g (0.02%)
Orange G loading dye (10x)	Ficoll 400 SDS (20% solution) EDTA Orange G Qs to 20ml dH_2O	3g (15%) 50µl (0.05%) 800µl (20mM) 0.125%
PBS-T (PBS-Tween)	Tween-20 in PBS (1x)	0.2% (v/v)
PBS-T (PBS-Triton)	Triton X-100 in PBS (1x)	0.2% (v/v)
PBS-T blocking buffer	Semi-skimmed milk (Marvel) Tween-20 in PBS (1x)	5% (w/v) 0.2% (v/v)
TBS-T (TBS-Tween)	Tween-20 in TBS (1x)	0.2% (v/v)
TBS-T (TBS-Triton)	Triton X-100 in TBS (1x)	0.2% (v/v)
TBS-T blocking buffer	Semi-skimmed milk (Marvel) Tween- 20 in TBS (1x)	5% (v/v) 0.2% (v/v)
1% Triton lysis buffer (store at -20°C)	Tris-HCl (1M, pH 8.0) EDTA (0.5M) Triton X-100 Qs to 50ml dH_2O Protease inhibitors (complete)	1ml (20mM) 1ml (10m) 1% (v/v) 1 tablet to 10ml

Buffer/ Solution/ loading dye	Component	Volume/ weight/ concentration
TBE buffer (20x)	Tris base Boric Acid EDTA (0.5M, pH 8.0) Qs to 1L dH_2O	1.78M 1.78M 40mM
Western blot running buffer	NuPAGE SDS Running buffer (20x) (Tris-Acetate) Qs to 1L dH_2O	50ml
Western blot transfer buffer	NuPAGE Transfer buffer (20x) Methanol Qs to 1L dH_2O	50ml 100ml

2.1.2. PCR and cDNA synthesis reagents

Table 2.2. Table of PCR and cDNA synthesis reagents.

Reagent	Supplier	Catalogue Number
Bovine Serum Albumin (BSA)	New England Biolabs	B9000S
Dimethyl Sulfoxide (DMSO)	Sigma	D4540
Deoxynucleoside Triphosphates (dNTPs)	Thermo Scientific	4303442
$MgCl_2$	Sigma	N8080130
$MgSO_4$	New England Biolabs	B1003S
Multiscribe Reverse Transcriptase	Thermo Scientific	4311235
PCR buffer	Thermo Scientific	N8080130
PCR enhancer	Thermo Scientific	11495017
Q5 High-Fidelity DNA polymerase	New England Biolabs	M0491S
Q5 reaction buffer	New England Biolabs	M0491S
Random Hexamers	Thermo Scientific	N808-0127
Rnase Inhibitor	Thermo Scientific	N8080119
Taq DNA polymerase	Sigma	D1806-1.5KU
Taq reaction buffer	Sigma	D1806-1.5KU

2.1.3. LUHMES cell culture reagents

Table 2.3. Table listing LUHMES cell culture reagents.

Reagent	Supplier	Catalogue Number
Advanced DMEM/F12	Life Technologies	12634028
b-FGF	Biolegend	571502
Geltrex LDEV-Free Reduced Growth Factor Basement Membrane Matrix	Thermo Scientific	A1413202
N6,2′-O-Dibutyryladenosine 3′,5′-cyclic monophosphate sodium salt (cAMP)	Sigma	D0627
Distilled Water	Life Technologies	15230089
DMSO	Sigma	D4540
Dulbeco PBS (DPBS)	Sigma	D8537
Fibronectin	Sigma	F1141
GDNF (recombinant, human)	R&D	212-GD-010
L-glutamine	Life Technologies	25030081
N2 supplement	Life Technologies	17502-048
Pierce 16% Formaldehyde (PFA) (w/v), Methanol-free	Thermo Scientific	28908
Poly-L-ornithine, hydrobromide (PLO)	Sigma	P3655
Tetracycline hydrochloride	Sigma	T7660
TrypLE Express Phenol Red	Life Technologies	12605010

2.1.4. Primary neuron culture reagent

Table 2.4. Table with primary neuron culture reagents.

Reagent	Supplier	Catalogue Number
B27 supplement	Thermo Scientific	17504-044
DMEM/F12 medium	Thermo Scientific	21331-020
DNAse I	Sigma	AMPD1
Dulbeco PBS (DPBS)	Sigma	D8537
EDTA	Sigma	E7889
Floxuridine	Sigma	50-91-9
GlutaMAX	Thermo Scientific	35050-061
Laminin	Thermo Scientific	23017-015
Leibovitz's L-15 Medium	Thermo Scientific	11415-064
Neurobasal Medium	Thermo Scieitific	21103-049
Papain	Worthington Biochemical	PAP2
Poly-D-Lysine hydrobromide	Sigma	P7886-50MG
Primocin	Lonza	VZA-1021
Uridine	Sigma	58-96-8

2.1.5. CRISPR/Cas9 reagents

Table 2.5. A table listing all reagents used in the CRISPR/Cas9 genome editing experiments.

Reagent	Supplier	Catalogue Number
ATP (10mM)	New England BioLabs	P0756S
DTT (10mM)	Thermo Scientific	R0862
***E.coli* Stbl3 strain**	Life Technologies	C7373-03
FastDigest BbsI	Thermo Scientific	FD1014
PlasmidSafe buffer (10x)	Epicentre	E3101K
PlasmidSafe exonuclease	Epicentre	E3101K
Tango buffer (10x)	Thermo Scientific	BY5
T4 ligation buffer (10x)	New England BioLabs	B0202S
T4 PNK	New England BioLabs	M0201S
T7 ligase with rapid ligation buffer (2x)	Enzymatics	L602L

2.1.6. Oligonucleotides used in the CRIPSR/Cas9 genome editing experiments

Table 2.6. Table listing all oligonucleotides used in the CRISPR/Cas9 genome editing experiments.

Oligonucleotide	Sequence
Oligonucleotides used to knock out SORCS2 in LUHMES cells	
***SORCS2* exon 1 top**	CACCGCGGGATGGCTTCGCGGGCGC
***SORCS2* exon 1 bottom**	AAACGCGCCCGCGAAGCCACTCCGC
***SORCS2* exon 3 top**	CACCGGTAGAAATTGTCGATGACGG
***SORCS2* exon 3 bottom**	AAACCCGTCATCGACAATTTCTACC
***SORCS2* exon 25 top**	CACCGACATCAGCTTCCGCTCCGAT
***SORCS2* exon 25 bottom**	AAACATCGGAGCGGAAGCTGATGTC
Oligonucleotides used to knock out *SORL1* in LUHMES cells	
***SORL1* exon 31 top**	CACCGTCGGTACCCGTCGCACACCC

***SORL1* exon 31 bottom**	AAACGGGTGTGCGACGGGTACCGAC
Oligonucleotides used to introduce the G508S mutation in LUHMES cells	
***SORL1* exon 10 top**	CACCGACTTACCAGTGGCGATGATG
***SORL1* exon 10 bottom**	AAACCATCATCGCCACTGGTAAGTC
ssODN	5’ CCCTTCATCTGGCTCAGCGCCTCAGTCAGCTCCTCAA CCTCCAGCTCCGGAGAATGC 3’

2.1.7. PCR and sequencing primers

Table 2.7. Table with PCR and sequencing primers.

Primer	Sequence
***SORCS2* exon 1 forward**	5’ CCTTTCTCTGCGCTCTCG 3’
***SORCS2* exon 1 reverse**	5’ CCGCCCCTGATGACCATA 3’
***SORCS2* exon 3 forward**	5’ CAGAGTGCCCAGGACTGTAC 3’
***SORCS2* exon 3 reverse**	5’ ATGTGCCCTAGGTATGCAGG 3’
***SORCS2* exon 25 forward**	5’ CCTCCACTGACAACCGCTTT 3’
***SORCS2* exon 25 reverse**	5’ GGCCCTTTCTAGATCTGCCT 3’
***SORL1* exon 10 forward**	5’ TTTACAGGGAACGCTAGGCA 3’
***SORL1* exon 10 reverse**	5’ TTGAGAACAGGCAAGCACAC 3’
***SORL1* exon 9-12 forward**	5’ CCATTTGCTGACTTCCACCG 3’
***SORL1* exon 9-12 reverse**	5’ TAGTGAGGTCCAGGAAGTGC 5’
***SORL1* exon 31 forward**	5’ CTGCTCAGAGCTGTGCCAGT 3’
***SORL1* exon 31 reverse**	5’ AGCCTTCCCTGGAGGTACAC 3’
pSpCas9(BB)-2A-GFP (Px458)	5’ TTTATGGCGAGGCGGCGG 3’

2.1.8. RT-PCR primers

Table 2.8. Table listing the RT-PCR primers used to screen for SORCS2 isoforms.

Primer	Sequence
***SORCS2* pair (1) forward**	5’ GATCAGCTTCCTCCTGCGAG 3’
***SORCS2* pair (1) reverse**	5’ CCTGGCCGTTTCCTTTTGAA 3’
***SORCS2* pair (2) forward**	5’ TTCAAAAGGAAACGGCCAGG 3’
***SORCS2* pair (2) reverse**	5’ AGTTGATGCTCAGGACCACG 3’
***SORCS2* pair (3) forward**	5’ CTGTCCAGGGCAACCACTCA 3’
***SORCS2* pair (3) reverse**	5’ CTCCGTGGGTGAAAAGACAGA 3’
***SORCS2* pair (4) forward**	5’ TCATCCTCTACAAGTTCAAAAGCAG 3’
***SORCS2* pair (4) reverse**	5’ CCTGAGTGGTTGCCTAGAGTCTG
***SORCS2* pair (5) forward**	5’ AGTTCAAAAGGAAACAGGCCAG 3’
***SORCS2* pair (5) reverse**	5’ GCCTGAGTGGTTGCCTAGAG 3’
***SORCS2* pair (6) forward**	5’ GATCCCCACGCAGGCAAC 3’
***SORCS2* pair (6) reverse**	5’ CCCTCCGTGGGTGAAAAGAC 3’
***SORCS2* pair (7) forward**	5’ TTCAAAAGCAGGAAACGGCCA 3’
***SORCS2* pair (7) reverse**	5’ CAGCGTCCACATGTCTCAGAT
***SORCS2* pair (8) forward**	5’ TGCTGAACGCACAGAAGATCA 3’
***SORCS2* pair (8) reverse**	5’ GCCTGGCCGTTTCCTGATGA 3’

2.1.9. qRT-PCR probes

Table 2.9. Table listing the qRT-PCR probes used.

Gene	Probe Number	Supplier
ATP5B	ge-DD-6-ATP5B	Primerdesign ltd.
ENOX2	ge-DD-6-ENOX2	Primerdesign ltd.
ERCC6	ge-DD-6-ERCC6	Primerdesign ltd.
RNF20	ge-DD-6-RNF20	Primerdesign ltd.
RPLPO	Hs00420895_gH	Applied Biosystems
SCLY	ge-DD-6-SCLY	Primerdesign ltd.
SDHA	ge-DD-6-SDHA	Primerdesign ltd.
SORCS1	Hs00364666_m1	Applied Biosystems
SORCS2	Hs00325181_m1	Applied Biosystems
SORCS3	Hs01039447_m1	Applied Biosystems
SORL1	Hs00268342_m1	Applied Biosystems
SORL1	Hs00983791_m1	Applied Biosystems
SORT1	Hs00361760_m1	Applied Biosystems
TBP	Hs00427620_m1	Applied Biosystems
UBC	ge-DD-6-UBC	Primerdesign ltd.
UBE4A	ge-DD-6-UBE4A	Primerdesign ltd.

2.1.10. Antibodies

Table 2.10. Table listing all antibodies used.

Antibody	Concentration	Supplier	Catalogue Number
Primary Antibodies			
53BP1	1:1000	Abcam	ab36823
β-tubulin	1:1000	Abcam	ab15568
γH2A.X	1:50 (IF); 1:400 (ICC)	Merk	05-636
GAPDH	1:10, 000	Merk	MAB374
SORLA	1: 750	R&D systems	AF5699
SORCS2	1:750	R&D systems	AF4238
SORCS2 (custom made)	1:1000	Dako	F7100
Vinculin	1:1000	Cell Signalling	13901S
Secondary Antibodies			
Alexa Fluor® 488 Donkey Anti-Mouse IgG (H+L)	1:500	Thermo Scientific	A21202
Alexa Fluor® 568 Donkey Anti-Rabbit IgG (H+L)	1:500	Thermo Scientific	A21207
Rabbit Anti-Mouse Immunoglobulins/HRP	1:1000	Dako	P0260
Rabbit Anti-Sheep Immunoglobulins/HRP	1:1000	Dako	P0163

2.2. Methods

2.2.1. Animals and behavioural experiments

2.2.1.1. Animals

Mice were bred and housed at the animal facility at Aarhus University, in groups of up to five mice per cage. WT and knockout mice were derived from different cohorts as breading was done on homozygous mice (i.e. full knockout and full WT breading). Mice of the same genotype were littermates; all mice (knockout and WT) were on the same background: C57Bl/6JBomTac (Taconic). Behavioural experiments were carried out using sex- and age-matched mice- male, 5-6 months old. Each of the behavioural tests described below were carried out in Denmark, using naïve animals in a randomized order by an investigator blinded to the mouse genotype. No animals were excluded from the subsequent analysis.

2.2.1.2. Exploration of a novel environment

Mice in the control group (here defined as 'home cage') were kept in their original cages. Mice in the novel environment ('Novel E') and the recovery from the novel environment ('recovery') groups were transferred to the testing room, where they were individually exposed to a novel environment for 2h. The novel environment consisted of an Open Field Arena with four different novel objects and mint-like odour. After the exploratory activity, the mice in the Novel E group were sacrificed. The mice belonging to the recovery group were returned to the initial cages, where they recovered from the behavioural task for 24h before being sacrificed together with the mice from the home cage group. The behaviour of the mice during the novel environment task, as well as in their home cages was not recorded.

2.2.1.3. Perfusion and tissue processing

Mice were perfused transcardially with cold PBS containing heparin (10,000 U/L), followed by ice-cold 4% paraformaldehyde (PFA) in phosphate-buffered saline (PBS). Whole brains were dissected and post-fixed overnight in 4% PFA in PBS. Following post-fixation, brains were rinsed in sterile PBS, transferred to fresh PBS and shipped to Edinburgh. Here, they were cryoprotected first in 10% sucrose and then in 30% sucrose at 4^{o}C until the tissue sank to the bottom of the tube. Brains were subsequently embedded in OCT compound on dry ice and stored at -80^{o}C. Coronal sections (14µm thick) containing the brain areas of interest (i.e. dentate gyrus, the

CA2 and the CA3 region of the hippocampus) were obtained using the cryostat and mounted directly onto Superfrost slides (Cat No 10149870, Thermo Scientific). Slide were stored at -80^{o}C until required. For the initial experiments, three serial sections corresponding to bregma -1.64mm were mounted onto the same slide and stained as described in 2.2.4. For the subsequent experiments, three brain sections corresponding to bregmas -1.64mm, -2.155mm and -2.555mm were obtained, mounted onto the same slide and stained.

2.2.2. LUHMES cell culture protocol

Low passage number proliferating LUHMES cells (P3) were kindly provided by Dr Faith Davies and Professor Cathy Abbott (Institute of Genetics and Molecular Medicine, The University of Edinburgh).

LUHMES cells were grown and differentiated at 37°C in a humidified 95% air and 5% CO2 atmosphere. Cell culture flasks were pre-coated with PLO (1mg/ml) and fibronectin (1mg/ml) in distilled H_2O (dH_2O) for at least 3 hours at 37^{o}C. Following incubation, the coating solution was aspirated, and the flasks washed two times with dH_2O and air-dried before cell seeding.

LUHMES cells were thawed by warming up the cryovial in a water batch set at 37^{o}C. 5ml of warm Advanced DMEM/F12 without additives were slowly added to the thawed cells, which were then centrifuged at 300g for 5 minutes. The supernatant was aspirated, cells resuspended in 5ml warm proliferation medium and transferred to a pre-coated T25 cell culture flask. Proliferation medium consisted of Advanced DMEM/F12, L-glutamine (200 mM), N2 supplement (100x), and b-FGF (160 µg/ml). Proliferating LUHMES cells were maintained until reaching about 85% confluence and then passaged. Medium was aspirated and cells were detached by adding trypLE, collected in Advanced DMEM/F12 without additives and centrifuged at 300g for 5 minutes at room temperature (RT). The cell pellet was resuspended in proliferation medium. For generating frozen stocks, cells were resuspended in proliferation media containing 20% foetal calf serum (FCS) and 10% DMSO and transferred to pre-cooled cryovials. Cryovials were stored in Mr Frosty™ for at least 48 hours at -80°C before being transferred for storage in liquid nitrogen.

Proliferating LUHMES cells were differentiated according to the protocol described by Scholtz *et al.* (2011)[133]. Four million cells were plated in a pre-coated T75 flask

containing 12ml of proliferation medium. After 24 hours, proliferation medium was exchanged with differentiation medium, consisting of Advanced DMEM/F12, L-glutamine (200 mM), N2 supplement (100x), cAMP (100mM), tetracycline (1 mg/ml), and recombinant human GDNF (20 µg/ml). Following 48 hours of predifferentiation, cells were detached as described above, counted and replated in pre-coated cell culture plates. Seeding densities and differentiation medium volume are listed in Table 2.11. Differentiated cells were grown up to day 14 and fed every second day by replacing half the medium with freshly prepared differentiation medium.

Table 2.11. Table showing the seeding densities and differentiation medium volume required for a given size of cell culture plastic plate.

Cell culture plate	**6 well**	**12 well**	**24 well**
Growth area (cm^2)	9.6	3.5	1.9
Volume of medium per well	2ml	1ml	500µl
Cell number to seed per well	$1.1x10^6$	$0.55x10^6$	$0.15x10^6$

2.2.3. Primary hippocampal neuron culture

Acid-etched coverslips, placed in 24-well plates, were coated with Poly-D-lysine for at least 2 hours at 37°C. Following incubation, the coverslips were washed three times with dH_2O and air-dried for 30 minutes. Laminin was then added to each coverslip and incubated for another 2 hours at 37°C.

Postnatal day 0 (P0) WT and *Sorcs2*$^{-/-}$ mouse pups were decapitated, the brains removed, and hippocampi isolated in ice cold DPBS. The dissected tissue was immediately transferred to ice cold Leibovitz's L-15 medium and kept on ice. Subsequently, the hippocampi were centrifuged for a minute at 1200rp, the supernatant was carefully removed and the sedimented tissue was digested for 30 minutes at 37°C in filter-sterilised papain solution. The papain solution was prepared in advance and contained 10ml of pre-warmed L-15 medium, 40µl EDTA (0.5M) and 167µl papain. The pH of the solution was adjusted by adding 60µl of NaOH (0.2M). After 30 minutes, the digestion was stopped by adding 5ml of DMEM/FCS (i.e. DMEM supplemented with 10% FCS, primocin (100µg/ml) and 5µl of DNAse I. The tissue was centrifuged for 1 minute at 1200rpm and resuspended until fully dissolved in 1ml

of DMEM/FCS containing DNAse I (0.01mg/ml). Further 10ml of DMEM/FCS were added and the suspension was centrifuged for 5 minutes at 1200rpm. The supernatant was removed carefully, and the pellet was resuspended in Neurobasal medium containing B-27 Supplement (1ml of B27 in 50ml of Neurobasal medium), GlutaMAX (2mM), Primocin (100 µg/mL), floxuridine (20µM) and uridine (20µM). Cells were counted and seeded at a density of 150,000 cells per coverslip. The neurons were grown for 12 days at 37°C, in 5% CO_2. Cells were fed every second day with Neurobasal medium supplemented with B27 before being fixed with 4% PFA in PBS.

2.2.4. Immunofluorescence on frozen brain sections

A modified immunofluorescence version of the protocol described by Crowe *et al.* (2011)[116] was performed as follows: brain sections containing areas of interest were thawed at RT for 15 minutes, incubated for 10 minutes in 4% PFA in PBS and then thoroughly washed in PBS (1X) containing 100 mM glycine (Cat. No 3570, EMD Millipore) followed by PBS (1X), each wash for 30 minutes. Antigen retrieval was subsequently performed by placing the slides in sodium citrate buffer (1X), pH 6.0 (Cat. No C1909, Sigma) heated to 95°C, and pulse-heated for 20 minutes in the citrate buffer in the microwave. Slides were allowed to cool for 20 minutes inside the microwave, followed by 30 minutes at RT. Slides were then washed 3 times in PBS (1X) for 15 minutes each wash and incubated for 1.5 hours in blocking solution consisting of 5% normal goat serum (Cat. No D9663, Sigma), 1% BSA (Cat. No 422351S, VWR), 0.1% Triton-X and 0.05% Tween-20 in PBS (1X) at RT. Slices were subsequently incubated with mouse anti-γH2A.X primary antibody (1:50) diluted in PBS (1X) containing 5% normal donkey serum and 1% BSA overnight at 4°C. The day after, slides were further incubated for another 30 minutes at 37°C. Following three washes with PBS (15 minutes each), sections were incubated with 3% Sudan black solution prepared in 70% ethanol for 10 minutes at RT. Slides were rinsed in dH_2O and secondary antibody (Alexa fluor-donkey anti-mouse 488), diluted 1:500 in 5% normal goat serum in PBS (1x), was subsequently added. Sections were incubated with secondary antibody for 1 hour at 37°C, and washed 3 times in PBS (1x), followed by 3 times in dH_2O, each wash for 15 minutes. DAPI (1:1000 in PBS) was subsequently applied for 10 minutes and washed off with PBS (3 washes, 5 minutes each). Sections were mounted in ProLong Gold antifade mountant (Cat No. P36934, Thermo Scientific).

Specific areas within the dentate gyrus, the CA2 and the CA3 region of the hippocampus were selected. Z-stacked confocal images, with a step size of 0.25µm, were acquired on a Nikon STORM/ A1+ microscope at 60x magnification, using the NIS Elements software. The optimal laser intensity and gain, which did not give a signal in the no-primary antibody control, were established and kept constant among all experimental groups. Three images, belonging to the frontal, middle and dorsal part of the hippocampus, were obtained for each mouse. Images were blinded and analysed using the software package Fiji[143]. The number of neurons having one or more γH2A.X-positive foci, as well as the total number of nuclei within a given area were manually counted. The percentage of γ-H2A.X-positive nuclei was determined for every image.

2.2.5. Creating mutant cell lines using CRISPR/Cas9 genome editing

2.2.5.1. Cloning of gRNA oligonucleotides into the pSpCas9(BB)-2A-GFP (Px458) vector

Guide RNAs (gRNAs) (Table 2.1.) were received as powder and lyophilised in dH_2O to a final concentration of 100µM. gRNAs were phosphorylated and annealed as follows.

The following master mix was prepared for each gRNA:

Component	**Volume (µl)**
gRNA top (100µM)	1
gRNA bottom (100µM)	1
T4 ligation buffer, 10x	1
T4 polynucleotide kinase	1
ddH_2O	6

The mixture was then incubated in a thermocycler using the following parameters: 37°C for 30 minutes, followed by 95°C for 5 minutes and ramped down to 25°C at 5°C per minute.

The phosphorylated and annealed gRNAs were subsequently diluted 1:200 by adding 1µl of the gRNA mix to 199µl of ddH_2O. The gRNAs were then ligated into the pSpCas9(BB)-2A-GFP plasmid (kindly provided by Dr Lora Boteva and Professor

Nick Gilbert, Institute of Genetics and Molecular Medicine, The University of Edinburgh) by setting up the following reaction:

Component	Volume (µl)
Diluted gRNA duplex	1
pSpCas9(BB)-2-GFP, 100ng	1
Tango buffer, 10x	2
DTT, 10mM	1
ATP, 10mM	1
Fast Digest BbsI	1
T4 ligase	0.5
ddH_2O	11.5

The ligation reaction was incubated for a total of 1 hour, comprising of 6 cycles of 37°C for 5 minutes, followed by 21°C for 5 minutes. Following incubation, the reactions were cleaned from any residual linearised DNA by treating them with PlasmidSafe ATP-dependent DNase for 30 minutes at 37°C, followed by 70°C for 30 minutes.

Component	Volume (µl)
Ligation reaction	11
PlasmidSafe Buffer	1.5
ATP, 10mM	1.5
PlasmidSafe endonuclease	1

The reaction was then stored at -20°C until transformation.

2.2.5.2. gRNA transformation

The pSpCas9(BB)-2A-GFP plasmid, with the gRNA ligated into it, was transformed into One Shot Stbl3 chemically competent *E.coli* (Cat No C737303, Thermo Scientific) as follows: A vial of Stbl3 cells was thawed on ice for each transformation reaction. 2µl of the PlasmidSafe-treated ligation reaction was added to the cells and mixed gently. 1µl of the pUC19 control (10pg) was included for estimating the transformation efficiency. The cell-DNA mixes were incubated on ice for 30 minutes, and then heat-shocked at 42°C for 45 seconds, followed by 2 minutes on ice. 250µl pre-warmed (37°C) S.O.C. medium was added to each vial, after which the vials were incubated at 37°C for 1 hour with constant horizontal agitation. Following incubation, three

different volumes (25µl, 75µl and 150µl) were spread on pre-warmed ampicillin-selective L-broth (LB) agar plates and incubated at 37^{o}C overnight.

The next day single colonies were identified, and ten colonies per gRNA were individually picked into 5ml LB, containing 100µg/ml ampicillin. After an overnight incubation at 37^{o}C with constant agitation, glycerol stocks were prepared by adding 200µl of 50% glycerol to 800µl of each culture. The mix was transferred to a cryovial and stored at -80^{o}C. DNA was isolated from the rest of the culture using a mini prep kit (Cat No. 27104, Qiagen), following the manufacturer's instructions. The obtained DNA was then sequenced to insure insertion of the gRNA into the pSpCas9(BB)-2A-GFP plasmid.

Once the correct gRNA insertion was confirmed using sequencing analysis, the desired plasmids were recovered from the glycerol stocks by scraping some bacteria and streaking it onto a pre-warmed LB agar plate, containing ampicillin (100µg/ml). Plates were incubated at 37^{o}C overnight. As before, single colonies were picked initially into 5ml of LB with ampicillin and incubated for 6 hours at 37^{o}C, which were then transferred to 250ml of LB with ampicillin (100µg/ml) and incubated overnight at 37^{o}C. DNA was extracted using the Endo-free Maxiprep kit (Cat No. 12362, Qiagen) following the manufacturer's instructions.

2.2.5.3. Transfection of proliferating LUHMES cells

Low passage number proliferating LUHMES cells (P5) were grown as described above until reaching 75% confluence. Two hours before transfection, cells were fed with fresh proliferation media. Subsequently, cells were harvested and counted. For each transfection reaction, 2x10^{6} cells were aliquoted into a new 4ml transfection tube and centrifuged at 90g for 10 minutes at RT. All medium was carefully aspirated, and the cell pellet was resuspended in 100µl freshly prepared nucleofect solution provided by the basic nucleofector kit for primary neurons (Cat No VAPI-1003, Lonza). The required DNA was added to the mixture. Generally, 2µg of the Cas9 plasmid containing the gRNA of interest was transfected, and in the experiments that required the incorporation of a repair template- 10µl of ssODN at concentration 10 µM. A transfection reaction where 2µg of a GFP-positive control plasmid was added to the cells served as a positive control. Each suspension was then transferred to a nucleofector cuvette and nucleofected using program D-33 on the Amaxa Nucleofector II D device. 500µl of prewarmed RPMI (Cat No BE12-752F, Lonza) was

gently added to each cuvette following nucleofection. The cells were then incubated at 37^{o}C for 5 minutes and gently added to precoated 6-well plates containing 2ml of freshly made proliferation medium.

Forty-eight hours following transfection, cells were lifted as described before and centrifuged at 90g for 10min. The cell pellets were then resuspended in 500µl of warm DPBS. Cells that were successfully transfected were sorted into precoated 96-well plates, containing 100µl of freshly prepared proliferation medium, by Elisabeth Freyer (Institute of Genetics and Molecular Medicine, The University of Edinburgh) using FACS (fluorescence-activated cell sorting). After seven days, 100µl of fresh proliferation medium was added to each well, and after ten days- single cell colonies were identified and transferred to 24-well plates for further expansion. In general, at this stage, one third of the cells was kept for genotyping, and the rest were split into two wells of a 24-well plate.

2.2.6. Genotyping of the generated mutant clones

2.2.6.1. DNA extraction

The collected cells were centrifuged at maximum speed for 5 minutes at RT. The media was gently removed, and the cell pellet was resuspended in 200µl of PBS. DNA was isolated using the DNeasy Blood and Tissue kit (Cat No 69504, Qiagen), according to the manufacturer's instructions. The quality and concentration of the obtained DNA was assessed using NanoDrop.

2.2.6.2. General PCR protocol

PCR amplification was carried out using a Sigma protocol in a total reaction volume of 20µl as follows: 0.2µl Sigma Taq Polymerase (5U/µl), 2µl Sigma buffer (10x) and 0.6µl dTNPs (5mM), 0.5µl primers (20µM).

PCR reaction with additives was set up for the GC-rich regions using 0.3µl Sigma Taq Polymerase (5U/µl), 2µl Sigma buffer (10x), 2µl PCR enhancer, 0.52µl MgSO4 (50mM), 0.6µl BSA (1/10), 0.6µl DMSO, 0.6µl dNTPs (5mM) and 0.6µl primers (20µm).

PCR with high fidelity DNA polymerase was set up for genotyping the generated G508S knock-in clones as follows: for 25µl reaction- 0.25µl Q5 high-fidelity DNA polymerase, 5µl Q5 reaction buffer (5x), 0.5µl dNTPs (10mM), 2.5µl primers (10µM).

All used primer pairs are listed in Table 2.7. Template and water volumes were adjusted based on the starting genomic DNA concentration.

Generally, all PCR reactions were run on a standard touch-down PCR program, altering the Primer Melting Temperature.

Temperature	Time	
93°C	1min	
93°C	20secs	
65°C	30secs	x 10 cycles
minus 1°C per cycle for 10 cycles		
72°C	1 min	
93°C	20secs	
65°C	30secs	x 30 cycles
72°C	1 min	
72°C	10mins	

Hold at 10°C until ready to take off block

End

2.2.6.3. General sequencing protocol

PCR products were prepared for Sanger sequencing. Generally, 1µl of PCR product (10-50ng/ 100bp) was cleaned from excess nucleotides and primers by using ExoSap-IT (Cat No 78250.40, Life Technologies). The following master mix was prepared:

Component	Volume (µl)
PCR product	1
ExoSap-IT	1
dH_2O	3

The reactions were incubated at 37°C for 60 minutes, followed by 80°C for 20 minutes in a thermocycler.

The following sequencing reaction was set up for each sample, using the primers listed in 2.1.7.

Component	Volume (µl)
ExoSap-IT treated PCR product	5
Big Dye v3.1 (Cat No 4337457, Thermo Scientific)	1
Big Dye sequencing buffer, (5x, Thermo Scientific)	1
Primer (3.2µM)	1
dH_2O	2

The reactions were run in a thermocycler using the following conditions:

Temperature	Time	
96°C	1min	
96°C	10sec	
		x 30
50°C	5sec	cycles
60°C	4 min	
Hold at 10°C until ready to take off block		

The sequencing reactions were subsequently cleaned by adding 2.5µl of EDTA (125mM), followed by 30µl of 95% ethanol. The samples were mixed by inverting four times and incubated for 15 minutes at RT. Subsequently, they were centrifuged at 3000rpm for 30 minutes at 4°C, the ethanol was removed, and the DNA was washed from excess salts by adding 30µl 70% ethanol to each reaction. Samples were then centrifuged for another 15 minutes at the same speed, the ethanol was removed, and

the pellet was left to air dry at RT for 30 minutes. Reactions were stored at -20°C prior to processing on 3730 Genetic Analyser (Applied Biosystems).

2.2.6.4. TOPO cloning

When a PCR reaction showed two bands on the agarose gel or failed to give a clean sequencing tracing, TOPO TA cloning was performed according to the manufacturer's instruction (Cat No 450071, Life Technologies). In brief, 2µl of fresh PCR product was mixed with 2µl dH_2O, 1µl of salt solution and 1µl TOPO vector provided with the TOPO TA cloning kit. The reaction was mixed and incubated at room temperature for 30 minutes, before being transformed into chemically competent *E. coli* cells (*DH5α*) as described in section 2.2.3.2. Twenty colonies were picked per PCR reaction. Each colony was boiled in 50µl dH_2O for 10 minutes and then used as a PCR template. PCR amplification was performed following the general PCR protocol described in 2.2.4.2.

2.2.7. Quantitative reverse transcriptase PCR (qRT-PCR)

Cell pellets from proliferating or differentiated LUHMES cells were obtained and resuspended in RLT buffer from the RNeasy mini kit (Cat No 74104, Qiagen) with 10% (v/v) 2-mercaptoethanol. Once resuspended, the pellets were immediately frozen on dry ice and stored at -80^{o}C until required. Total RNA was extracted from each sample using the RNeasy mini kit, according to the manufacturer's instructions. The quality and concentration of the obtained RNA was assessed using the NanoDrop.

Total RNA (1µg per sample) was reverse transcribed with Multiscribe Reverse Transcriptase using random hexamers in a 80µl reaction, following the manufacturer's protocol. A 'minus RT' control, in which 25ng RNA of each sample was used to make cDNA in the absence of the Multiscribe Reverse Transcriptase, was included to detect genomic contamination. Meanwhile, a 'no RNA' containing all components of the kit used except for the RNA was also included. cDNA was then stored at -20°C until required.

PCR amplification of the cDNA obtained for each sample was quantified using the TaqMan® Universal PCR Mix No AmpErase® UNG (Cat No 4324018, Life Technologies), and the threshold cycle (Ct) was determined using the Applied

Biosystems 7900HT Fast Real-Time PCR System and the corresponding SDS software. TaqMan probes were used for the detection of both the genes of interest (GOIs) and the reference genes (see section 2.1.9.). Eight reference gene controls were analysed in the sample set and the GeNorm software was used to identify the most stably expressed ones. A standard curve, generated from a dilution series, was run for each of the GeNorm selected stable reference genes, as well as the GOIs. The baseline and Ct values were determined for each gene and the GOI expression levels were calculated using the standard curve method for absolute quantification, where unknowns are compared to the generated standard curve and values are extrapolated. The GOI expression values were subsequently normalised to the geometric mean of the reference genes.

2.2.8. Western blot analysis

2.2.8.1. Protein extraction

Protein was normally extracted from three wells of a six-well plate, with each well containing $1.1x10^6$ cells. After aspiration of cell medium, cells were placed on ice and washed once with ice cold DPBS. 50ul of ice cold 1% Triton lysis buffer were added to each well, and the plate was left on ice. After one minute, cells were scraped off the bottom of the well with a cell scraper, and the lysed cells from the three wells were collected into a single microcentrifuge tube. The lysates were then thoroughly mixed on a tube rotator at 4°C for 30 minutes, followed by a centrifugation at a maximum speed (13,300g) at 4°C for 15 minutes. Supernatants were transferred to a new tube and the pellets were discarded. 10µl of the lysate obtained were aliquoted into a new tube, and equal volume of Laemmli buffer (2x) was added to the rest of the sample. The former was used to quantify the protein concentration using the Bio-Rad BSA protein assay (Cat No 5000116, Bio-Rad).

2.2.8.2. Quantification of protein lysate concentration

The Bio-Rad BCA protein assay kit was used to quantify protein levels in cell lysate samples. Protein standards, ranging from 0mg/ml to 2mg/ml, were prepared by diluting a known concentration of BSA (2mg/ml, Cat No 23209, Life Technologies) with dH_2O as diluent (see below). 5ul of each protein standard concentration was

pipetted in duplicate into a 96 well plate. Cell lysates were diluted 1:2, and 5ul of each was pipetted into the plate in triplicate.

Table 2.12. Table showing the amount of BSA (2mg/ml) and dH_2O required for generating a given concentration from the standard curve.

Concentration (mg/ml)	**2mg/ml BSA (µl)**	**dH_2O (µl)**
0	0	20.0
0.2	2.0	18.0
0.4	4.0	16.0
0.6	6.0	14.0
0.8	8.0	12.0
1.0	10.0	10.0
1.2	12.0	8.0
1.4	14.0	6.0
1.6	16.0	4.0
1.8	18.0	2.0
2.0	20.0	0

The assay working reagent was made up by mixing reagents A and S in the ratio 50:1. 25ul of reagent A+S was added to each well, followed by 200µl of reagent B. The plate was incubated at RT for 10 minutes, and absorbance was measured at 450nm using a FLUOstar Omega plate reader. Concentrations of cell lysates were extrapolated from the standard curve.

Following quantification, Laemmli sample buffer (2x) was added in a 1:1 ratio, and the samples were subsequently boiled in a heat block for 5 minutes at 95^{o}C. Samples were stored at -20^{o}C until required.

2.2.8.3. SDS-PAGE electrophoresis

NuPAGE Tris-acetate (3 -8%) precast gels were purchased from Life Technologies (Cat No EA03752BOX) and placed in a XCell SureLock Mini-cell tank according to the manufacturer's instructions. Approximately 200ml of Tris-acetate running buffer (Cat No LA0041, Life Technologies) were added to fill the upper (inner) buffer

chamber, with the buffer level exceeding the level of the wells. Up to 25µl of each sample (based on the estimated total protein concentration) were added to the wells, alongside 7µl of All Blue protein ladder (Cat No 1610373, Bio-Rad). The lower (outer) chamber was filled in with running buffer, and the gel was run at 150V for 1 hour and 30 minutes until the blue dye front reached the bottom of the gel.

2.2.8.4. Western blot transfer

After electrophoresis, proteins were transferred from the gel to a membrane using the wet blotting method. Transfer buffer was prepared by making up NuPAGE Transfer buffer (1x) in dH_2O with 10% methanol (or 20% when large proteins were examined). Pads were subsequently soaked in transfer buffer until saturated and any air bubbles were removed by squeezing them out before use. PVDF membranes (Cat No LC2002, Life Technologies) were activated in methanol for 30 seconds, briefly washed in dH_2O and equilibrated in transfer buffer for several minutes.

The tank was disassembled, and the gel was removed from the plastic cassette using a gel knife. A transfer sandwich was assembled by placing a piece of pre-soaked filter paper on top of the gel, after which the gel was turned over and the pre-wet membrane was placed on the gel, followed by another pre-soaked filter paper. Any bubbles were removed using a plastic roller. Three pre-soaked blotting pads were placed in the cathode core of the transfer module. The transfer sandwich was carefully placed on them, followed by another three pre-wet blotting pads. The anode core was placed on top of the pads, and the assembly was transferred to the XCell SureLock Mini-cell tank and secured in place with a gel tension wedge provided. Transfer buffer was added to the inner chamber, and 500ml of dH_2O were added to the lower chamber. Transfer was carried out at 30V for 1.5 hours at 4^oC

2.2.8.5. Immunostaining using HRP-conjugated secondary antibodies

Following transfer, the membrane was blocked in 5% dried semi-skimmed milk in TBS-0.2% Tween 20 for 1 hour at RT with constant agitation. Primary antibodies (listed in section 2.1.10) were diluted in blocking solution and added to the membrane. The membrane was incubated with the antibody solution at 4^oC overnight.

The next day, the membrane was washed three times for 10 minutes in TBS-0.2% Tween 20 and subsequently incubated for 1 hour at RT with secondary HRP-conjugated antibody (raised against the host species of the primary antibody used) diluted in 5% dried semi-skimmed milk in TBS-0.2%Tween 20. After another three

washes with TBS-0.2%Tween 20, blots were visualised using the Pierce ECL Plus Western Blotting Substrate (Cat No 11527271, Thermo Scientific), according to the manufacturer's instructions. The blots were wrapped in cling-film and exposed using autoradiography film.

2.2.9. Immunocytochemistry

Pre-differentiated LUHMES cells were plated down ($0.15x10^6$ cells per well) and grown on acid-etched coverslips, placed in 24-well plates and coated with PLO and fibronectin, followed by geltrex. After 14 days of differentiation, neurons were fixed with 4% PFA in PBS for 10 minutes at RT, after which they were washed once in warm PBS and stored in TBS at 4ºC until required.

After a quick wash with TBS, cells were permeabilised by adding 500µl of 0.2% Triton in TBS (1x) to each well for 5 minutes. Following three quick rinses with TBS, cells were incubated in blocking solution consisting of 5% normal donkey serum in 0.2% Tween-TBS for 1 hour at RT. Mouse monoclonal anti-γH2A.X primary antibody (1:400) and rabbit polyclonal anti-53BP1 primary antibody (1:1000) were diluted in blocking solution and added to the coverslips. Following an overnight incubation at 4ºC, the cells were washed with 0.1% Tween-TBS (3x10 minutes) and then incubated with the corresponding antibodies for 1 hour at RT. Secondary antibodies were Alexa Fluor 488-donkey anti-mouse IgG (1:300) and Alexa Fluor 596-donkey anti-rabbit IgG (1:500), and were diluted, together with DAPI (1:1000), in 4% normal donkey serum in 0.1% Tween-TBS. Cells were then washed with TBS (3x 10 minutes) and mounted with ProLong Gold antifade mountant (Cat No P10144,Thermo Scientific).

Z-stacked confocal images, with a step size of 1µm, were acquired on a Nikon STORM/ A1+ microscope at 100x magnification, using the NIS Elements software. The optimal laser intensity and gain, which did not give a signal in the no-primary antibody control, were established and kept constant among all experimental groups. Imaged areas were chosen based on the DAPI staining to avoid bias.

Once obtained, the images were blinded and analysed using the software package Fiji[143]. Around a hundred nuclei (corresponding to four images from different areas within the same coverslips) were analysed for each cell line. Nuclei with apoptotic morphology, as examined based on the DAPI staining (i.e. condensed staining and nuclear 'blebbing') were excluded from the analyses. Nuclei completely covered with

intense γH2A.X staining were also excluded from the analysis, as it has been previously suggested that this staining pattern corresponds to apoptotic cells, as well. The total number of nuclei, as well as the number of foci positive for γH2A.X and 53BP1 were determined for each image. The average number of foci per nucleus was determined for each image, and the average of the four images was calculated for each clone.

2.2.10. ELISA

At day 2, pre-differentiated LUHMES neurons were plated down in 6-well plates, pre-coated with PLO and fibronectin, at density of $1.1x10^6$ cell per well. Neurons were fed every two days, starting from day 6, when 700µl of fresh differentiation media was added to each well. From there on, 900µl were taken out and 1ml of freshly made differentiation media was added to each well every other day. Following 14 days of differentiation, for each cell line medium from three wells was collected in a 15ml falcon tube and spun at 2700rpm for 20 minutes at 4°C. After spinning the supernatant was carefully collected, medium aliquoted and stored at -80°C until required. Cell lysates from the corresponding wells were collected as described above (section 2.2.6.1.). All steps were performed on ice.

$A\beta_{42}$ sandwich ELISA (Cat No. KHB3441, Thermo Scientific) and Dopamine competitive ELISA (Cat No. AMS.EA1301Hu, Amsbio) were performed following the manufacturer's instructions. Coefficient of variation (CV) was calculated for the technical replicates. When the CV between the technical replicates was higher than the intra-assay variation (CV<3.6% for the $A\beta_{42}$ ELISA and CV<10% for the Dopamine ELISA), the most different value from the replicate was discarded. The average of the replicates for each sample was calculated and the concentration of extracellular $A\beta_{42}$/ Dopamine was extrapolated from the standard curve. The obtained values were normalised to the protein concentration (as calculated using the BSA assay) in the respective cell lysates from the same wells.

2.2.11. Alamar Blue Assay

Pre-differentiated LUHMES neurons (day 2) were plated down in 24-well plates, pre-coated with PLO and fibronectin, at 150,000 cells/well. Neurons were fed every two days, starting from day 6, when 300µl of fresh differentiation media was added to each

well. From there on, 250µl were taken out and 300µl of freshly made differentiation media was added to each well every other day.

Viability was measured at day 6 and day 14 by replacing the medium with freshly made differentiation medium containing 10% (v/v) Alamar Blue (Cat No DAL1025, Thermo Scientific). Cells were incubated with the Alamar Blue solution for 2 hours. After 2 hours, the solution was transferred to a new 24-well plate and fluorescence measured in a FLUOstar OMEGA plate reader using an excitation wavelength of 540-570nm, and an emission wavelength of 580-610nm.

2.2.12. Statistical analysis

For the 'novel environment' experiment' the results obtained for the *Sorcs2*$^{-/-}$, *Sorl1*$^{-/-}$ and WT mice were analysed using the R package 'nparLD'. NparLD was specifically developed for repeated measurement data in factorial experiments. It utilises rank-based, nonparametric, as such, is for small datasets. The data was first stratified into independent groups based on the experimental group (i.e. 'home cage', 'novel E' and 'recovery') and the genotype (i.e. WT and KO mice). The repeated measurements for each mouse comprised of images taken from the frontal, middle and dorsal part of the hippocampus (so individual mice constituted the repeated grouping factor).

The rank mean percentage of γH2A.X- positive nuclei in each of the three experimental groups ('home cage', 'novel E' and 'recovery') within the same genotype (WT or KO mice) was compared for each brain region (dentate gyrus, CA2 or CA3 region of the hippocampus) using the "F1.LD.F1" function of the R package "nparLD". Two homogenous groups (environments) were measured repeatedly (brain areas) in x experimental units (mice). This was coded as '*Y (percentage of γH2A.X- positive nuclei) = repeated measure (brain area) + group (environment) + repeated measure (mouse)*'. The same function was used to compare the rank mean percentage of γH2A.X- positive nuclei between two genotypes (WT and KO mice) for each experimental group ('home cage', 'novel E' and 'recovery') and brain region (dentate gyrus, CA2 or CA3 region of the hippocampus). Two homogenous groups (genotypes), measured repeatedly (brain areas) in x experimental units (mice). This was coded as '*Y (percentage of γH2A.X- positive nuclei) = repeated measure (brain area) + group (genotype) + repeated measure (mouse)*'. The F2-LD-F1 model was used to compare the slopes (i.e. the rate of break generation and recovery) between

two genotypes for each brain region (dentate gyrus, CA2 or CA3 region of the hippocampus). In this case four homogenous groups (two different genotypes within two environments) were measured repeatedly (brain areas) in x experimental units (mice). This was coded as '*Y (percentage of γH2A.X- positive nuclei) = repeated measure (brain area) + group (genotype) + group (environment) + repeated measure (mouse)*'. The mean rank differences between grouping factors were tested using the non-parametric ANOVA-type statistic for all comparisons.

In all other analysis, GraphPad Prism version 7.00 for Windows (GraphPad Software, La Jolla California USA, www.graphpad.com) was used throughout to draw graphs and perform unpaired student's t-tests.

Sample size estimation and post-hoc statistical power analyses were calculated using the GPower 3 software, taking into account the corresponding effect sizes, α error probability (set to 0.05), the total sample size and the number of groups. Power was defined as '1- β error probability'. Experiments where the power was below 0.8 were considered underpowered. Where non-parametric tests were performed, power and sample size were approximated based on equivalent parametric tests. The sample size required for a non-parametric test was determined by multiplying the sample size calculated for the corresponding parametric test by a correction factor, referred to as asymptotic relative efficiency (ARE).

Chapter 3 DNA damage and repair in *Sorcs2*$^{-/-}$ and *Sorl1*$^{-/-}$ mice

3.1. Introduction

3.1.1. Background

The maintenance of cellular genome integrity is vital for long-lived cells, such as neurons where excessive and/or unrepaired damage will otherwise lead to neuronal loss. Exacerbated DNA damage has been implicated in the process of brain aging and age-related cognitive decline, as well as in neurodegenerative disorders including AD[144, 145, 146, 147].

Recently, experiments with mice have associated DSB formation and repair with learning and memory processes. Suberbielle *et al*. (2013)[107] showed that in WT mice, exploration of a novel environment triggers the formation of transient DSBs in brain areas important for learning and memory as a result of increased neuronal activity. Subsequent experiments have suggested that DSB formation, in the context of neuronal activity, is a physiological mechanism, which depends upon the activation of Topoisomerase IIβ and is required for the rapid expression of early-response genes[118].

Meanwhile, hAPP-J20 mice (an AD mouse model) not only showed increased DSB formation at baseline, but also increased levels of DNA damage following recovery from an exploratory activity. This suggests that excess Aβ levels increase DSB formation and impair DNA repair. Additionally, in WT primary hippocampal neurons, both Aβ- and activity-induced formation of DSBs relied on the activation of NR2B-containing NMDARs, while their repair on that of NR2A-containing NMDARs[107].

3.1.2. Aim

My aim was to investigate the impact of knocking out either *Sorcs2* or *Sorl1* in mice on activity induced DSB formation. I hypothesised that, bearing in mind the role of the sortilin family in APP processing[148] and processes relevant to learning and memory, knocking out *Sorl1* and/or *Sorcs2* in mice could potentially replicate the effects of elevated Aβ levels on both generation and repair of activity induced DSBs.

3.1.3. Experimental Design

To address this hypothesis, I assessed the levels of γH2A.X (a marker for DNA DSB formation) in *Sorcs2*$^{-/-}$ and *Sorl1*$^{-/-}$ and corresponding WT mice at baseline, as well as immediately after or following recovery from exploration of a novel environment.

I received two sets of mouse brains from our collaborators, Assoc. Prof Simon Glerup and Dr Ditte Olsen, Aarhus University, Denmark. The first set of brains was derived from *Sorcs2*$^{-/-}$ and corresponding WT mice, while the second- from *Sorl1*$^{-/-}$ and corresponding WT mice. The two sets of mice (*Sorcs2*$^{-/-}$ and *Sorl1*$^{-/-}$, and corresponding WTs) were bred independently from each other, and I received the brains at different times. All mice were on the same background: C57Bl/6JBomTac (Taconic). WT and knockout mice were derived from different cohorts as breading was done on homozygous mice (i.e. full knockout and full WT breading), and mice of the same genotype were littermates. All mice used in the experiments below were sex- and age-matched- male, 5-6-month-old, and WT and knockout mice from the same batch were bred simultaneously.

The behavioural part of the experiment was performed by Dr Ditte Olsen and Mathias Kaas Ollendorff in Denmark. In brief, mice in the control group (here defined as 'home cage') were kept in their original cages and sacrificed at the end of the experiment. Mice in the novel environment ('Novel E') and the recovery from the novel environment ('recovery') groups were exposed to a novel environment that constituted an Open Field Arena with four different novel objects and mint-like odour. The mice were allowed to explore the novel environment for two hours after which the mice in the Novel E group were sacrificed, while the mice in the recovery group were returned to their home cages, where they recovered from the behavioural task for 24 hours before being sacrificed (Figure 3.1.). The behaviour of the mice in the home cages, as well as during exploration of the novel environment was not recorded. However, it has been previously reported that the *Sorcs2*$^{-/-}$ mice exhibit increased activity in a novel environment, which disappears following habituation. No changes in the activity levels of *Sorcs2*$^{-/-}$ mice in the home cage has been observed (Ditte Olsen; manuscript in preparation).

In both sets, the mice were sacrificed by direct perfusion with 4% PFA through the circulatory system. After perfusion, the brains were dissected and shipped to

Edinburgh. I received whole brains on which I performed the experiments described below.

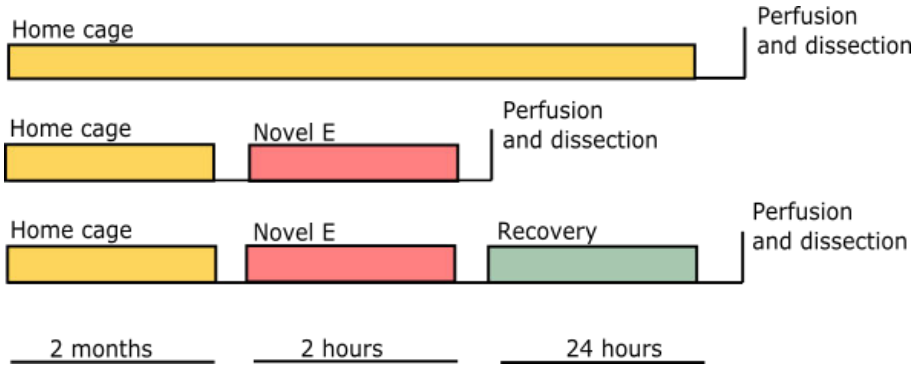

Figure 3.1. Experimental design of the behavioural experiments performed by our collaborators in Denmark. Mice from all genotypes (WT, Sorcs2$^{-/-}$ and Sorl1$^{-/-}$) were divided in three groups: 'home cage' (yellow), 'Novel E' (pink) and 'recovery' (green). The 'home cage' group shows the levels of DNA DSBs at a baseline, while the mice in the 'Novel E' and 'recovery groups'- the extent of DSB formation following exploratory activity and recovery from it, respectively.

3.2. Optimisation and validation of the γH2A.X immunostaining protocol

In order to assess the levels of DNA damage and repair in knockout and corresponding WT mice, I first developed an immunofluorescent (IF) protocol for staining frozen brain sections with γH2A.X.

Initially, I tested the IF protocol described by Suberbielle *et al.* (2013)[107] on frozen sections from a WT mouse brain, provided by Dr Matthieu Vermeren and Dr Toby Hurd (Institute for Genetics and Molecular Medicine, The University of Edinburgh). Using this protocol, I observed high levels of background staining even in the absence of both primary and secondary antibody, suggestive of auto-fluorescence (Figure 3.2.A). Thereafter, different approaches for optimising the IF protocol were undertaken, including blocking with different solutions (such as BSA and/or donkey or goat serum), increasing the blocking and washing times and using different antibody (both primary and secondary) concentrations. Following more than ten trial experiments, some progress in reducing the background auto-fluorescence was made (Figure 3.2.B), but its generally high levels, combined with the lack of detectable

nuclear staining, suggested that either the γH2A.X antibody or the used IF protocol did not work.

In order to determine whether or not the used γH2A.X antibody was specific for detecting DNA DSBs in tissues, frozen mouse testes (provided by Dr James Crichton and Dr Ian Adams, Institute for Genetics and Molecular Medicine, The University of Edinburgh) were used as a positive control, since phosphorylation of the H2A.X histone has been shown to mark main events of the spermatogenic process associated with chromatin reorganisation[149]. After an extensive literature review, I tested six IF protocols, previously used for staining frozen testis sections and spermatogenic cell cultures (modified in order to be used in tissue) with γH2A.X. Following this process, I developed a protocol that detected nuclear staining in the testis sections (Figure 3.2.C). Subsequently, I tested this protocol on sections from the WT mouse test brain. Unfortunately, when used on frozen brain tissue, no γH2A.X staining and high background levels were seen again (Figure 3.2.D).

Given that the γH2A.X antibody detected DNA DSBs in frozen testes, but not brain sections, it was concluded that either the protocols I had been trying so far were incompatible or the mouse brain I had been using had very low, undetectable levels of DSB formation. To account for the latter, I contacted Prof Tara Spires-Jones (Centre for Discovery Brain Sciences, The University of Edinburgh) who provided me with an APP/PS1 mouse brain. The APP/PS1 mouse is an AD mouse model similar to the hAPP-J20 mice used by Suberbielle *et al*.[107]. Thus, the APP/PS1 mice were likely to exert high levels of baseline neuronal DSB formation as the hAPP-J20 mice used by Suberbielle *et al*. Around the same time, I came across a study describing a γH2A.X IF protocol for staining frozen rat brain sections[116]. Using a slightly modified version of this protocol, I detected a faint γH2A.X signal in brain sections I obtained from the APP/PS1 mouse (Figure 3.2.E). Increasing the γH2A.X antibody concentration (from 1:400 to up to 1:10) improved the signal to noise ratio (Figure 3.2.E- H).

Subsequently, the staining was further optimised on brain sections from the beginning of the hippocampus (bregma: -1.28 mm) of *Sorcs2*$^{-/-}$ mice that belonged to the Novel E group, as it was hypothesised that this group was likely to show the highest level of DNA DSB formation. As for the APP/PS1 mouse, I was able to detect γH2A.X staining at 1:100 dilution (Figure 3.3.A). Even though I could detect the γH2A.X signal, the levels of auto-fluorescence were still quite high. In order to overcome this, I decided

to include a Sudan Black step in the protocol, as well as to use green- instead of red-fluorescent secondary antibody. Using a different secondary antibody and Sudan Black before, but not after, incubation with secondary antibody, reduced the levels of auto-fluorescence drastically. After titrating the primary antibody concentration, the 1:50 dilution was selected as the optimal one as it was the lowest showing the best signal to noise ratio (Figure 3.3.B-D).

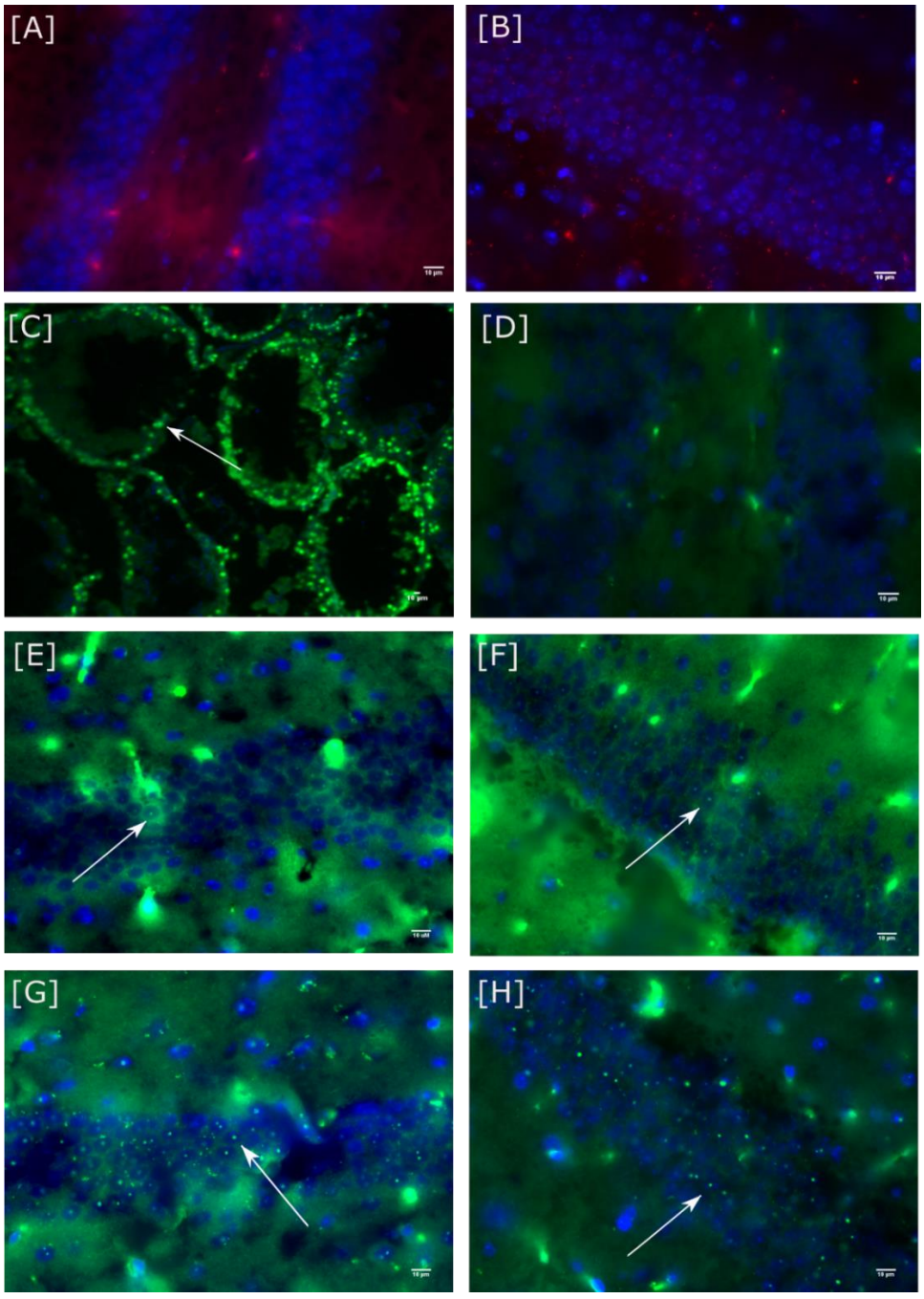

Figure 3.2. Optimisation of the γH2A.X IF protocol. [A-B] Representative images depicting autofluorescence levels (red) and the lack of γH2A.X staining (red) in the test WT mouse brain,

acquired with epifluorescent microscope at 40x magnification. [C] A representative image of mouse testis sections stained for γH2A.X (green) and DAPI (blue), obtained with epifluorescent microscope at 10x magnification. [D] A representative image of APP/PS1 mouse brain sections immunostained for γH2A.X (green) and DAPI (blue), following the protocol which gave the best result in the testes; image acquired with epifluorescent microscope at 40x magnification. [E-H] Representative images of APP/PS1 mouse brain sections immunostained for DAPI (blue) and γH2A.X (green) at concentrations 1:400 [E], 1:100 [F], 1:50 [G] and 1:10 [H], following the Crowe et al. (2011) protocol; images acquired with epifluorescent microscope at 40x magnification. In all images, white arrows point towards nuclei positive for γH2A.X, and scale bars correspond to 10um.

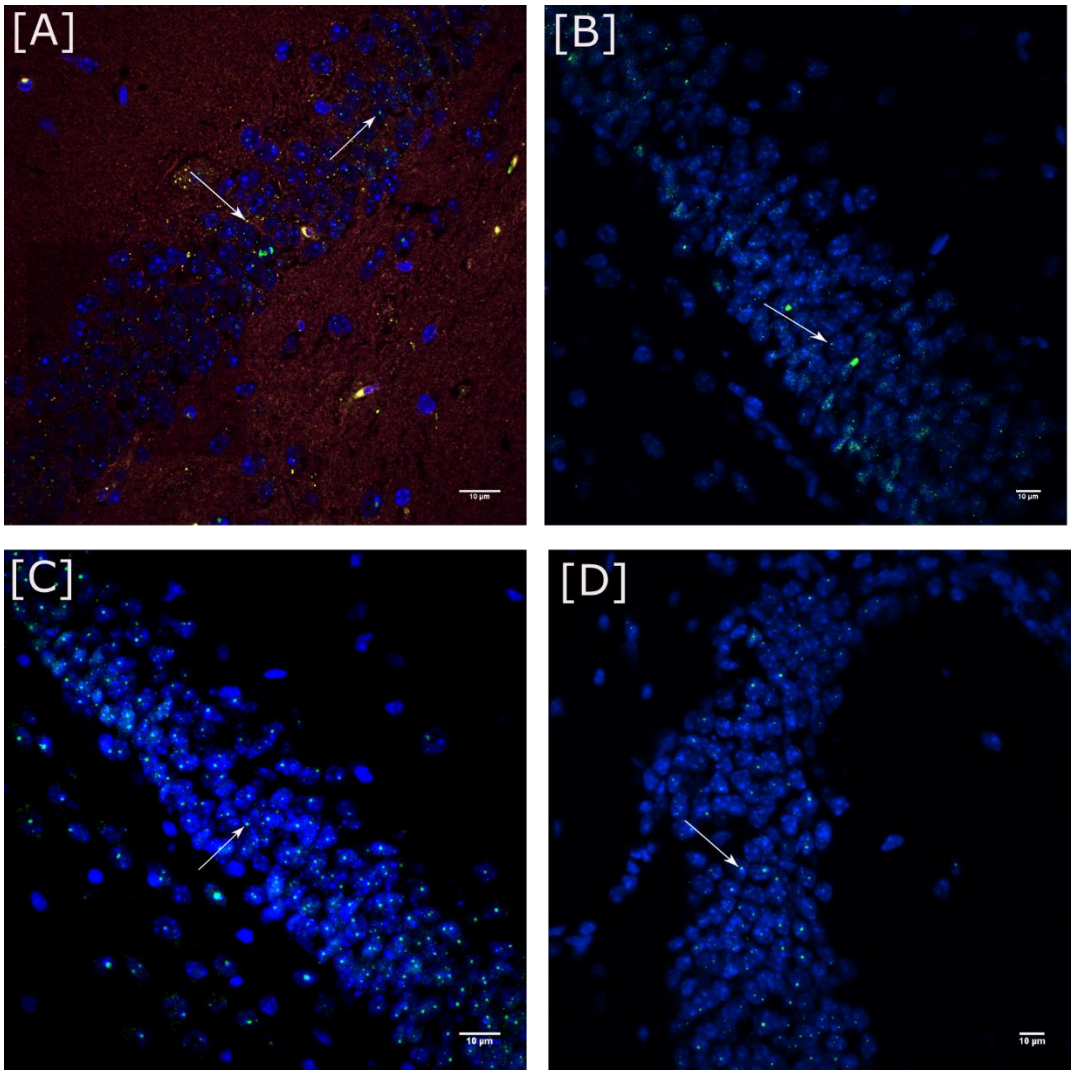

Figure 3.3. Further optimisation of γH2A.X IF protocol on Sorcs2$^{-/-}$ brain sections. [A] A representative image of Sorcs2$^{-/-}$ brain sections stained for DAPI (blue) and γH2A.X (green) at concentration 1:100, acquired with confocal microscope at 100x magnification. The red

channel depicts auto-fluorescence. White arrows point towards either γH2A.X foci (green) within nuclei or foci-like auto-fluorescence (yellow). [B-D] Representative images of Sorcs2$^{-/-}$ brain sections stained for DAPI (blue) and γH2A.X (green) at concentrations 1:100 [B], 1:50 [C] and 1:10 [D], following optimisation with Sudan Black. White arrows point towards nuclei positive for γH2A.X. The images were acquired using confocal microscope at 60x magnification. In all images, scale bars correspond to 10um.

3.3. An attempt of developing an automated method for assessing levels of DNA DSBs

In order to make the process of counting the total number of nuclei, as well as the percentage of nuclei positive for γH2A.X more efficient and unbiased, I tried to develop an automated method for assessing the levels of DNA DSBs using the software package Fiji[143], as it has been previously used for identifying γH2A.X foci within nuclei. This required the delineation (also known as 'segmentation') of the nuclei in order to subsequently use them as a 'mask' to identify γH2A.X foci within them. Unfortunately, due to the highly compacted distribution of the nuclei within the dentate gyrus and the CA regions of the hippocampus, precise nuclear segmentation was impossible (Figure 3.4.B). This inability to 'segment' and count the nuclei correctly prevented me from using this method of automated counting.

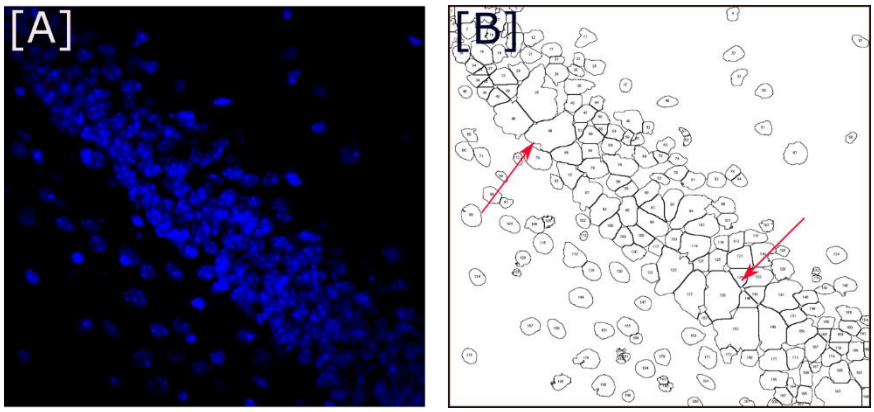

Figure 3.4. Attempt of developing an automated method for assessing DNA DSB formation. [A] A representative confocal image acquired for the dentate gyrus, showing only the nuclei in the DAPI channel (blue), 60x magnification. [B] At attempt to 'segment' and count the nuclei automatically using the Fiji software. Red arrows point at regions where the software was unable to separate the nuclei from each other correctly.

3.4. Preliminary results: assessing levels of DNA DSBs in *Sorcs2*$^{-/-}$ and *Sorl1*$^{-/-}$ mice

3.4.1. Assessing levels of DNA DSB formation in *Sorcs2*$^{-/-}$ and corresponding WT mice (preliminary results)

Initially, my goal was to identify the same area of the dentate gyrus and the CA3 region of the hippocampus for each animal and count the number of nuclei positive for γH2A.X within it. I stained three consecutive sections from the beginning of the hippocampus (bregma: -1.64 mm) of *Sorcs2*$^{-/-}$ and corresponding WT mice, belonging to one of the three experimental groups ('home cage', 'Novel E' and 'recovery') described in Figure 3.1. Using confocal microscopy, I took z-stacks from specific areas within the dentate gyrus (the end of the infra-pyramidal blade of the granule cell layer) and the CA3 hippocampal region. The exact location of the selected areas is indicated with a red rectangle in Figure 3.5.A. Within these areas, I counted manually the total number of nuclei, as well as those positive for γH2A.X, whilst blinded with respect to genotype and condition (Figure 3.5.B; Appendix A, Table A1).

One-way ANOVA identified a significant effect of the environment on the level of DNA DSB formation in the CA3 region of the WT mice ($p=0.0448$; Figure 3.5.C, left). However, post-hoc analysis did not identify any significant difference in the percentage of γH2A.X-positive nuclei between the three conditions ('home cage', 'Novel E' and 'recovery') in this brain region (Figure 3.5.C, left; Appendix Table A2). Meanwhile, I observed the same pattern as Suberbielle *et al.* (2013)[107], i.e. transient increase in the number of γH2A.X-positive nuclei following exploratory activity in the dentate gyrus of the WT mice (Figure 3.5.C, right). In this brain region, the percentage of γH2A.X-positive nuclei detected in the 'Novel E' group was significantly higher than the one detected in both the 'home cage' ($p=0.0007$) and the 'recovery' ($p=0.0033$) groups (Figure 3.5.C, right; Appendix Table A2). No difference in the levels of DSB formation was observed between the home cage and recovery groups ($p=0.2344$; Appendix Table A2).

Sorcs2$^{-/-}$ mice did not show the same pattern of a transient increase in the proportion of γH2A.X-positive nuclei following exploration of a novel environment in either the dentate gurus ($p=0.1050$) or the CA3 region ($p=0.4751$) of the hippocampus (Figure 3.5.D). Further inspection of the data (Appendix A, Table A1) suggested that these

differences were the result of increased DNA DSB formation in the *Sorcs2*$^{-/-}$ mice in the home cage group. Comparison of the percentage of γH2A.X-positive nuclei between knockout and wild type mice at baseline (home cage) indicated a significant difference in both the dentate gyrus (p= 0.0225) and the CA3 region (p= 0.0153) (Figure 3.5. E).

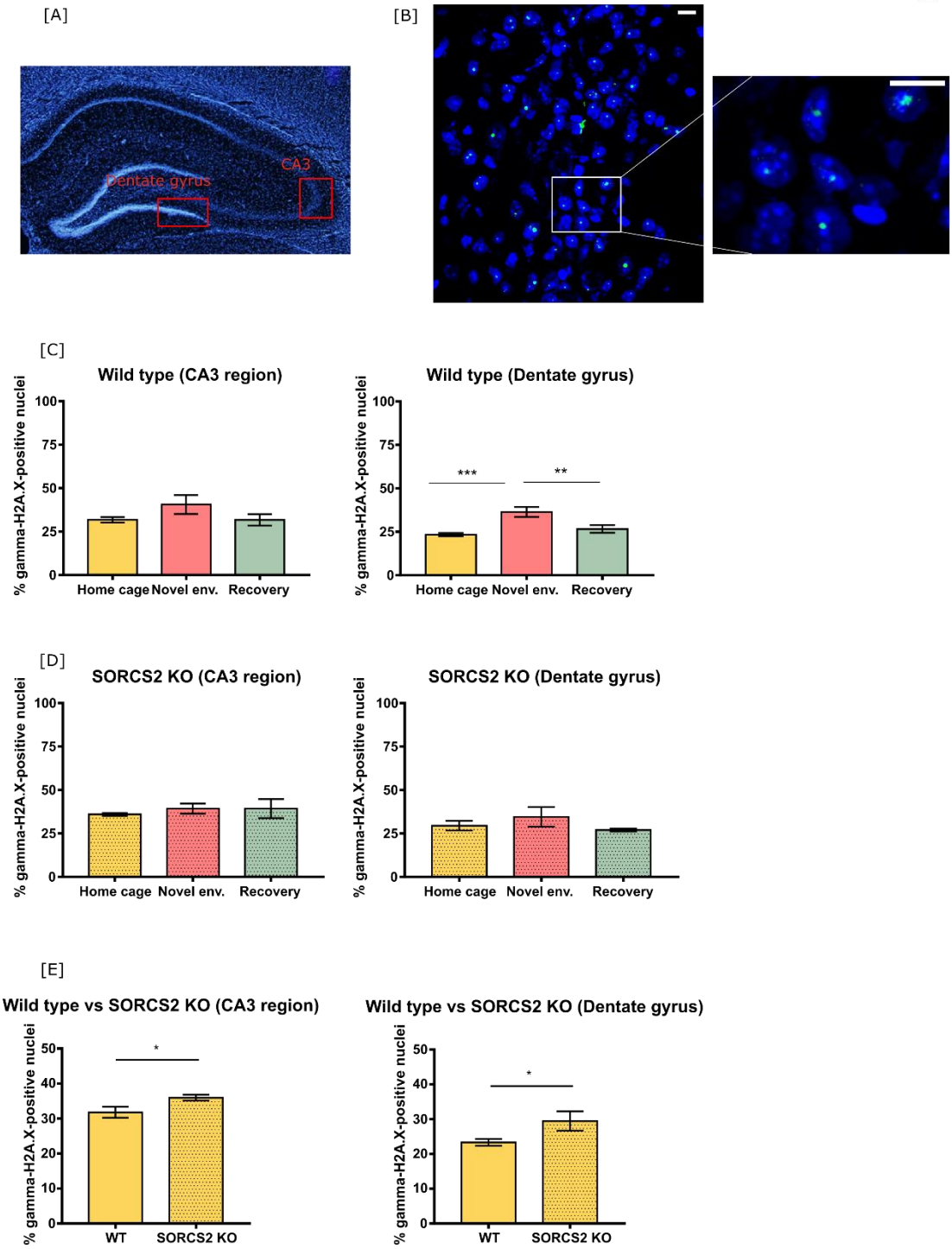

Figure 3.5. Assessing levels of DSB formation in the frontal part of the hippocampus of Sorcs2$^{-/-}$ mice and corresponding WT mice. [A] An image of the mouse hippocampus showing the regions of the dentate gyrus and the CA3, where the number of γ-H2A.X-positive nuclei was counted. Adapted after https://biodiscovery.eu/products/dapi/. [B] A representative image obtained with confocal microscope, showing DNA DSBs (stained with γH2A.X in green) within nuclei (stained with DAPI in blue), 60x magnification. Scale bars correspond to 10µm. [C-D] Percentages of γH2A.X-positive nuclei in dentate gyrus and the CA3 region of the hippocampus (averaged from three consecutive sections per mouse, n=3) in WT [C] and Sorcs2$^{-/-}$ [D] mice belonging to one of the three experimental groups. [E] Comparison between the percentage of γH2A.X-positive nuclei in the dentate gyrus and the CA3 region of the hippocampus of WT and Sorcs2$^{-/-}$ mice in the home cage group. Error bars represent means ± SD; One-way ANOVA [C-D] or unpaired t-test [E]; '', '**' and '***' indicate p < 0.05, p<0.001, p<0.001, respectively.*

3.4.2. Assessing levels of DNA DSB formation in *Sorl1*$^{-/-}$ and corresponding WT mice

As for the first batch of *Sorcs2*$^{-/-}$ and WT mice, I stained three consecutive sections from the beginning of the hippocampus (bregma: -1.64mm) of *Sorl1*$^{-/-}$ and corresponding WT mice, belonging to one of the three experimental groups described in Figure 3.1. Using confocal microscopy, I acquired z-stacks from one, specifically defined area within the dentate gyrus and the CA3 region of the hippocampus (as described in 3.4.1. and shown in Figure 3.5.A). Due to the long imaging time, half of the images were acquired by Dr Jonathan Phillips and the other half by me, using the same imaging parameters. Independent from each other and blinded to genotype and condition, we counted the total number and the number of γH2A.X-positive nuclei in all acquired images. The obtained counts are reported in Appendix A, Table A3.

No significant difference in the percentage of γH2A.X-positive nuclei was observed between the three conditions ('home cage', 'Novel E' and 'recovery') for either of the two genotypes in both the dentate gyrus (WT mice p=0.4859; *Sorl1*$^{-/-}$ mice p=0.2357) and the CA3 region of the hippocampus (WT mice p=0.6388; *Sorl1*$^{-/-}$ mice p=0.6571) (Figure 3.6.A-B; Appendix A, Table A4). Surprisingly, I observed that the WT mice in the 'home cage' group showed higher number of γH2A.X-positive nuclei than the WT mice from the *Sorcs2*$^{-/-}$ mice experiment (Appendix A, Table A1 and A3). Since the WT mice from the two batches were on the same background, I compared the

percentage of DSB-positive nuclei between the two groups.This indicated a significantly higher percentage of γH2A.X-positive nuclei in both the dentate gyrus (p=0.0498) and the CA3 region (p=0.0002) of the hippocampus of the WT mice that accompanied the *Sorl1*$^{-/-}$ mice compared to those that accompanied the *Sorcs2*$^{-/-}$ mice (Figure 3.6.C). Increased levels of DSBs were also observed in the CA3 region (p=0.0064), but not the dentate gyrus (p=0.8129) of *Sorl1*$^{-/-}$ mice compared to *Sorcs2*$^{-/-}$ mice (Appendix B, Figure B1).

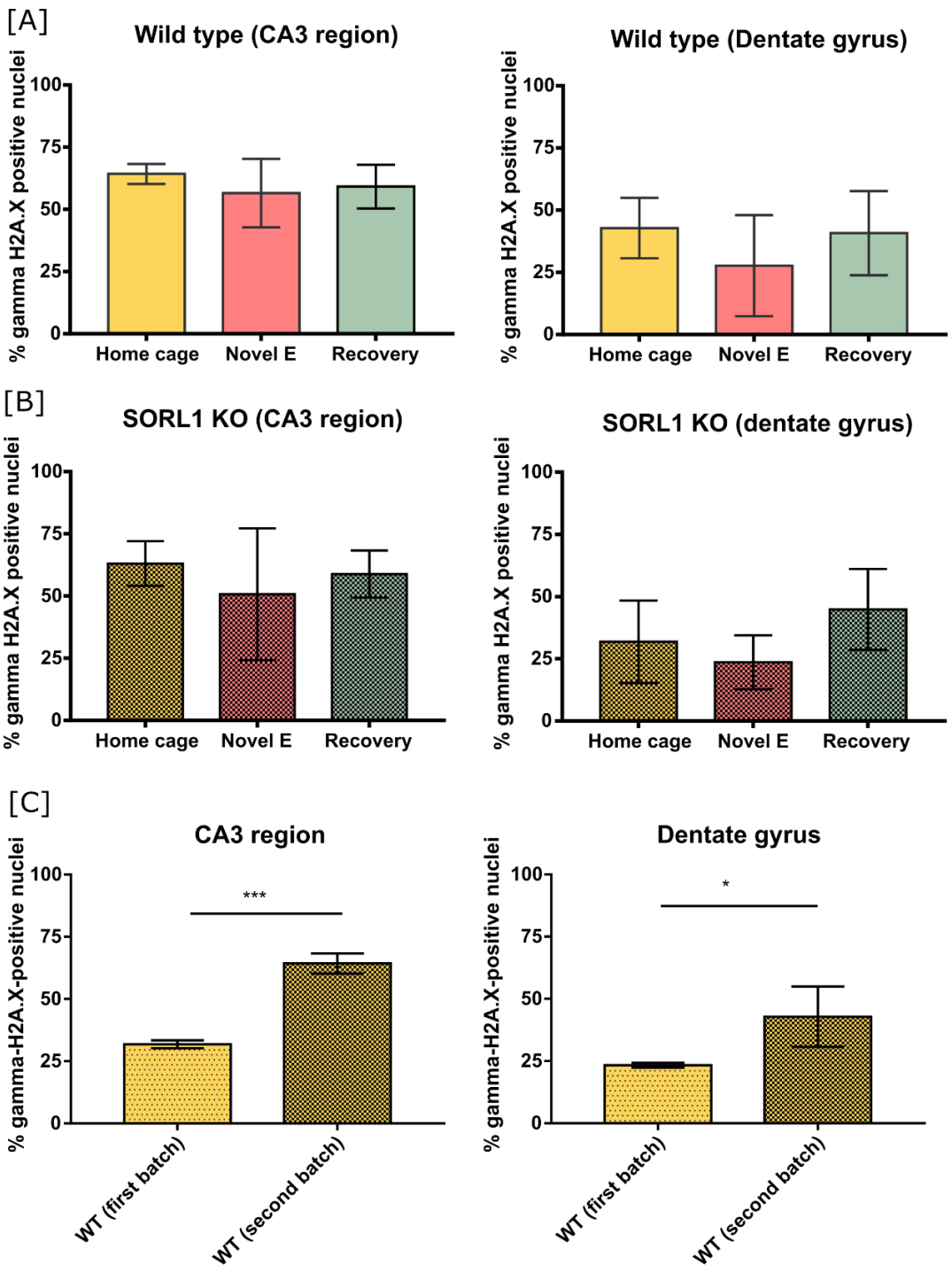

Figure 3.6. Assessing DSB formation in Sorl1$^{-/-}$ and corresponding WT mice. Percentages of γH2A.X-positive nuclei in dentate gyrus and the CA3 region of the hippocampus (averaged from three consecutive sections per mouse, n=3-4) in WT [A] and Sorl1$^{-/-}$ [B] mice, belonging to one of the three experimental groups. [C] Comparison between the percentage of γH2A.X-positive nuclei in the CA3 region of the hippocampus (left) and the dentate gyrus (right) of WT mice belonging to either the first or the second batch we received. Error bars represent means

*± SD; One-way ANOVA [A-B] or unpaired t-test [C]; '***' and '*' indicate p<0.001 and p < 0.05, respectively.*

3.5. Assessing levels of DNA DSBs in the *Sorcs2*$^{-/-}$ mice- sampling throughout the entire hippocampus

Following detection of a significant difference between the three experimental groups in the dentate gyrus of the WT mice that accompanied the *Sorcs2*$^{-/-}$ mice (Figure 3.5.C), as well as in the baseline DSB levels in the frontal part of the dentate gyrus and the CA3 region of the hippocampus of the assessed *Sorcs2*$^{-/-}$ mice (Figure 3.5.E), I went on to investigate if these differences were observed more widely throughout the hippocampus. This was suggested by the external expert in my thesis committee- Prof Tara Spires-Jones (Centre for Discovery Brain Sciences, The University of Edinburgh). This time, I decided to include the CA2 region of the hippocampus, as well, as *Sorcs2* has been previously shown to be highly expressed in that region[20]. I stained slices from the frontal (bregma: -1.64mm), the middle (bregma: -2.12mm) and dorsal end (bregma: - 2.75mm) of the hippocampus of *Sorcs2*$^{-/-}$ and WT mice, belonging to the described above groups. As described in 3.4.1, I obtained confocal images from specifically defined areas within the dentate gyrus and the CA2 and CA3 region of the hippocampus in these sections. As before, due to time constraints, half of the images were acquired by Dr Jonathan Phillips and the other half by me. Blinded to both genotype and condition, I counted the total number of nuclei and the number of nuclei positive for DSBs in the dentate gyrus and the CA2 and CA3 region of the hippocampus. In order to confirm that my counting was accurate and unbiased, 10% of the images (selected at random) were counted by Dr Phillips, who was also blinded to genotype and condition. Comparison of the counts showed 80-90% agreement between the two of us. Actual counts are reported in Appendix A, Table A5.

Following a discussion with Dr Mairead Bermingham (a post-doc in our group) and given the small sample size of my study (n=3 mice per genotype per experimental group), I decided to analyse the obtained results using 'NparLD'[150]. NparLD is an R-based package that was developed for analysing repeated measurement data in factorial experiments. In addition, the fact that it is non-parametric is helpful with small sample sizes, as this analysis does not require the data to be normally distributed. All

the analysis described below was performed by Dr Bermingham. The data was first stratified into independent groups based on the experimental group (i.e. 'home cage', 'novel E' and 'recovery') and the genotype (i.e. WT and *Sorcs2*$^{-/-}$), but also into repeated measurements (i.e. images taken from the frontal, middle and dorsal end of the hippocampus) with individual mice constituting the grouping factor. For each individual group (e.g. *Sorcs2*$^{-/-}$ mice belonging to the 'home cage' group), the percentage numbers of γH2A.X- positive nuclei were converted to ranks. The mean rank differences in the percentage of γH2A.X-positive nuclei were compared across the groups.

Using this method, as expected given previously published findings[107], a significant increase in the percentage of γH2A.X-positive nuclei was detected in the dentate gyrus ($p=4.53x10^{-9}$), the CA2 ($p=1.46x10^{-5}$) and the CA3 ($p=0.035$) regions of the hippocampus of WT mice following exploration of the novel environment (, Figure 3.7.A; Appendix A, Table A6). The percentage of γH2A.X-positive nuclei returned back to baseline levels following 24-hour recovery period in all three brain regions (min $p=1.78x10^{-6}$; Figure 3.7.A; Appendix A, Table A6). By contrast, no significant differences were found in the percentage of DSBs across the three experimental groups in any of the brain regions examined in the *Sorcs2*$^{-/-}$ mice (min $p=0.074$; Figure 3.7.B; Appendix A, Table A6). Moreover, as for the preliminary results described in 3.4.1, *Sorcs2*$^{-/-}$ mice showed statistically significantly higher percentage of γH2A.X-positive nuclei in the home cage group when compared to corresponding WT mice (Figure 3.8.A). However, this difference was observed only in the dentate gyrus ($p=0.025$), and not in the CA2 ($p=0.138$) or the CA3 region ($p=0.176$) of the hippocampus (Figure 3.8.A).

Further analysis also carried out by Dr Bermingham, compared the impacts of the three environmental conditions on the formation of DSBs between the genotypes. These models identified a significant difference in the way WT mice responded to the novel environment when compared to *Sorcs2*$^{-/-}$ mice. In the dentate gyrus, the percentage of γH2A.X-positive nuclei was significantly affected by the interaction between genotype and exposure to a novel environment ($p=0.001$, Figure 3.8.B, left). It appeared that, in addition to the greater number of DSBs observed in the home cage group, the *Sorcs2*$^{-/-}$ mice also do not acquire as many DSBs on exposure to the novel environment. Similarly, a significant genotype by environment interaction was observed in the CA2 ($p=0.021$) and CA3 ($p=0.049$) regions (Appendix B, Figure B2).

No significant interaction was identified between the genotype and the recovery from the novel environment (min p=0.139, Figure 3.8.B, right; Appendix A, Table A7; Appendix B, Figure B2). In these analyses- fitting the interaction between genotype and environment, the main effect of the novel environment was significant in the dentate gyrus and CA2, but not in the CA3 region (Appendix A, Table A8). No main effect of genotype was observed in any of the three brain regions examined (Appendix A, Table A9).

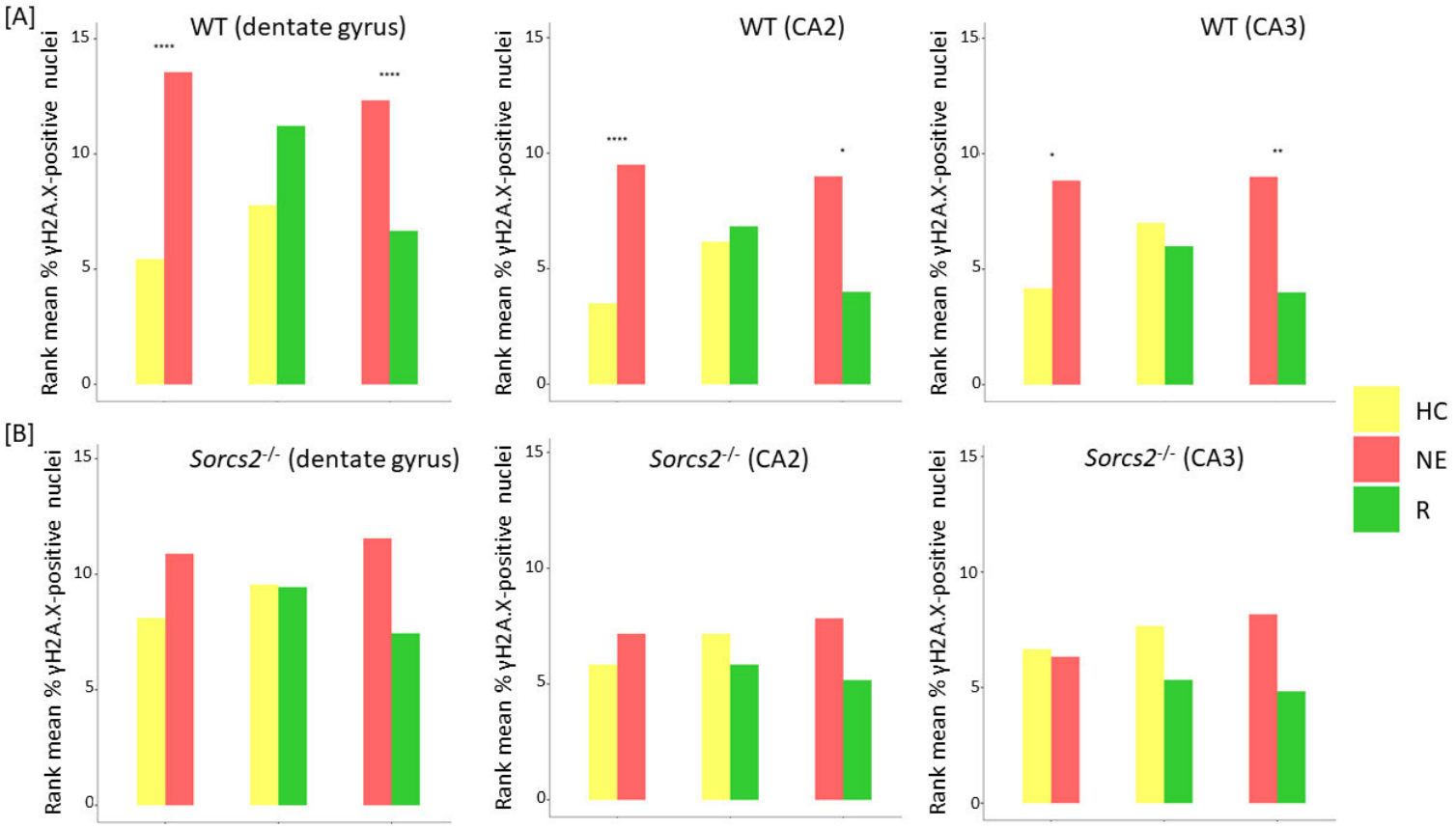

Figure 3.7. Comparison of the levels of DSB formation between the three experimental groups in WT and Sorcs2$^{-/-}$ mice. Rank mean percentages of γH2A.X-positive nuclei in dentate gyrus, the CA2 and CA3 region of the hippocampus in WT [A] and Sorcs2$^{-/-}$ [B] mice belonging to one of the three experimental groups-'home cage' (HC), 'Novel E' (NE) and 'recovery' (R).. N=3, non-parametric ANOVA type statistic, '', '**' and '****' indicate p<0.05, p<0.01 and p<0.0001, respectively;*

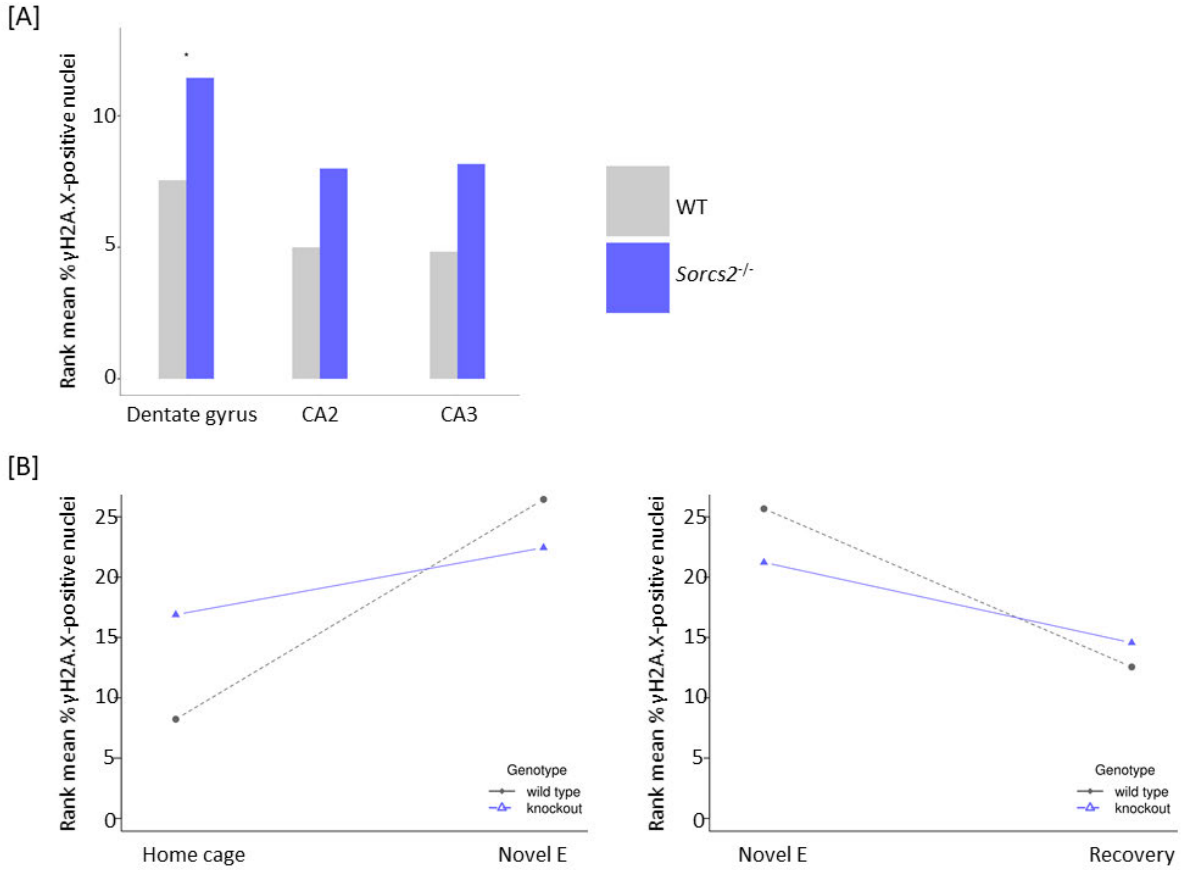

Figure 3.8. Assessing the difference in the levels of DNA DSB formation between WT and Sorcs2$^{-/-}$ mice belonging to one of the three experimental groups. [A] Comparison of the rank mean percentage of γH2A.X-positive nuclei in the dentate gyrus, the CA2 and CA3 region of the hippocampus of WT and Sorcs2$^{-/-}$ mice. [B] Comparison of the rate of DSB formation (left) and repair (right) between WT and Sorcs2$^{-/-}$ mice. N=3, non-parametric ANOVA-type statistic, '' indicates p<0.05.*

3.6. Discussion

As described above, I optimised and validated an immunofluorescence protocol for staining frozen brain sections using γH2A.X as a marker of DSB formation. Subsequently, I reproduced the previously reported finding that exploration of a novel environment triggers transient increase in DSB formation in the hippocampus of adult WT mice. Meanwhile, *Sorcs2*$^{-/-}$ mice from the same batch did not show the same pattern. They appeared to acquire fewer DSBs following the novel environment task compared to wild types, despite the fact that *Sorcs2*$^{-/-}$ mice have been previously reported to be more active in an open field test[6]. At the same time, *Sorcs2*$^{-/-}$ mice belonging to the 'home cage' group showed statistically higher levels of DSBs in the dentate gyrus.

A major caveat to the findings of this chapter is the relatively small number of animals used (n=3 for the *Sorcs2*$^{-/-}$ batch and n=3-4 for the *Sorl1*$^{-/-}$ batch), as well as the fact that each experiment was performed only once. Post-hoc power analyses revealed that both experiments were powered to detect changes in the formation of DSBs following exposure to the Novel E or recovery from it in the *Sorcs2*$^{-/-}$, *Sorl1*$^{-/-}$ and corresponding WT mice. These analyses were based on the results obtained for the WT mice accompanying the *Sorcs2* knockouts. For this group of mice, the differences in the proportion of γH2A.X-positive nuclei observed across the three conditions in all three brain regions were in the range of the minimal effect size (1.36) that the study was sensitive to detect for this sample size (n=3) and if powered to 80%. Even more so, the effect size observed when the level of DSB formation was compared across the three conditions in the dentate gyrus of WT mice was much higher- 2.88. Thus, the non-significant change in the number of γH2A.X-positive nuclei across the three experimental groups observed in the *Sorcs2*$^{-/-}$ mice is likely to represent a true lack of difference.

Similarly, these experiments were powered to assess the effect of the genotype-environment interaction, as well as of the environment alone on the formation of γH2A.X foci. In the case of the dentate gyrus, for example, the observed effect sizes for the genotype-environment interaction (1.17) and the environment alone (2.00) were much higher than the predicted by the sensitivity test change of 0.83 at 80% power. This experiment was, however, under-powered for assessing the effect of the genotype in this model. In the dentate gyrus, the predicted minimum genotype effect size (0.72) was bigger than the actual (0.50). Subsequent sample size analysis

estimated that a total of 42 animals (or 7 mice per genotype per condition) is required to reliably assess the effect of the genotype given the observed in this model effect size.

Meanwhile, despite being powered to detect differences in the number of DSB-positive nuclei across the three experimental groups, I did not observe the same pattern of DSB formation upon exploration of a novel environment in the WT mice that matched the *Sorl1*$^{-/-}$ mice. These WT mice displayed significantly higher levels of baseline DSBs in the dentate gyrus and the CA3 region of the hippocampus when compared to the WT mice from the first set. Given this unexpected finding, it was not possible to compare the levels of DSB formation in the *Sorl1*$^{-/-}$ mice either at baseline or following exploration of a novel environment. There are a number of possible reasons that can explain these surprising results. Perhaps the most likely explanation is changes in the animal facility that took place prior to the experiment. This was subsequently noted by our collaborators Dr Ditte Olsen and Mathias Kaas Ollendorff (personal correspondence). There were also some discrepancies in the way the novel environment experiment was performed in the two batches of mice. In the first experiment, WT and *Sorcs2*$^{-/-}$ mice were exposed to the novel environment individually. Subsequently, *Sorl1*$^{-/-}$ and matched WT mice performed the exploratory task together with their cage-mates. It has been previously reported that exposing mice to a novel environment with other, familiar mice reduces anxiety levels and increases exploratory activity[107]. Thus, it can be speculated that the elevated levels of DSBs I observed in the WT mice from the first batch upon exploratory activity might be due to increased stress levels from being exposed to novelty on their own, rather than the novel environment itself. However, this seems unlikely to be the case. Suberbielle *et al.* (2013)[107] showed that the observed increase in neuronal DSB formation following exploratory activity is solely dependent on neuronal activity, and not the release of corticosterone or other stress factors, as removing the adrenal gland exerted no effect on this process. However, in that study, mice were subjected to a novel environment with their cage-mates and with mice from a different litter. This might constitute an additional factor of novelty that has not been accounted for in the experiment described here.

These observations raise the question whether the novel environment protocol used in our study is sufficient to trigger the formation of DSB. Recently, Alvarez-Castelao *et al.* (2017)[151] described a new cage with an enriched novel environment (EE), which

comprised a labyrinth with changeable side walls, running wheels and toys of different size, texture and colour. Exposure to this EE altered the expression of over 200 proteins important for neuron and synapse function[151]. Although this study was performed in younger mice (six weeks of age) and entailed chronic exposure to the described EE (21 days), it would be interesting to look at the impact of this type of cage on the mice described here. In addition, both locomotor activity and habituation have been previously shown to decrease with the time WT mice spend in a novel environment. Therefore, two hours might be too long exposure period, enough for the mice to get used to the novel environment and stop exploring. Moreover, such habituation might subsequently allow DSB repair. Therefore, it would be also interesting to test whether shorter exposure to the described novel environment could result in a stronger and more reliable phenotype in WT mice.

Despite the caveats described above, the increased DSB formation observed in the *Sorcs2*$^{-/-}$ mice at baseline constitutes an interesting finding. It is unlikely that this could be explained by the hyperactivity phenotype in an open field test previously reported for these mice[6], since no differences in their behaviour or activity levels compared to the WT mice was observed for the home cage condition (Dr Ditte Olsen, personal correspondence). However, there are a few possible reasons of why the *Sorcs2*$^{-/-}$ mice might have higher levels of neuronal DSBs at baseline.

Hypothesis 1: Knocking out *Sorcs2* leads to elevated Aβ levels and thus DSB formation.

Given the link described above between Aβ and DSB formation, as well as that between members of the sortilin family and APP processing[148], an obvious hypothesis is that lack of SorCS2 in mice can lead to increased Aβ production and thus DSB formation. Increased Aβ levels have been reported in mice where *Sorl1*[37] or *Sort1*[152] have been knocked out, and knock-down of *SORCS2* in HEK293 cells led to increased APP processing[148]. However, to date, there is no study investigating the effect of knocking out *SORCS2* on Aβ accumulation in either mice or human neurons.

Hypothesis 2: Knocking out *Sorcs2* disturbs the balance between NR2A- and NR2B-containing NMDA receptors, leading to aberrant network activity and increased DSB formation.

As mentioned previously, the activation of NR2B- and NR2A-containing NMDA receptors has been shown to be necessary for the formation and repair of DSBs,

respectively. NR2A- and NR2B-containing NMDA receptors have been implicated in the induction of LTP and LTD, respectively[153]. Meanwhile, SorCS2 has been shown to be crucial for the induction of NMDAR-dependent hippocampal LTP and LTD in mice[5]. Additionally, SorCS2 has been associated with the recycling and trafficking of NR2A-containing NMDA receptors to dendritic and synaptic compartments of striatal medium spiny neurons[51]. Therefore, it could be hypothesised that lack of SorCS2 might result in impaired NMDAR function and consequently increased formation and/or insufficient repair of activity-induced DSBs. Since no DNA DSB repair deficits were observed in the *Sorcs2*$^{-/-}$ mice, it is more likely that knocking out *Sorcs2* impairs the formation rather than the repair of DSBs.

Hypothesis 3: Knocking out *Sorcs2* leads to increased oxidative stress and thus DNA damage

Oxidative stress has been shown to cause DSBs[154] and activation of NR2B-containing NMDA receptors has been linked to elevated levels of oxidative stress[155]. However, according to Suberbielle *et al.* (2013), treatment with a ROS scavenger had no effect on NDMAR-induced DSB formation. Moreover, no oxidative stress has been reported in hAPP-J20 mice at the age used in their study[107]. On the other hand, SorCS2 has recently been implicated in neuronal protection from oxidative stress by playing indirect role in the synthesis of the endogenous ROS scavenger glutathione. Lack of SorCS2 has been associated with increased oxidative damage and death of hippocampal neurons following PTZ kindling in a mouse model of epilepsy[22]. Therefore, it is reasonable to speculate that the increased baseline levels of DSB we detect in the *Sorcs2*$^{-/-}$ mice could be, at least partly, due to increased oxidative stress. However, Malik *et al.* (2019)[22] did not report increased baseline levels of oxidative stress in the *Sorcs2*$^{-/-}$ mice, but rather an exacerbated response to it.

In summary, increased Aβ levels have been linked to both increased oxidative stress[156] and altered balance between NR2A- and NR2B-containing NMDAR activity[157]. Likewise, enhanced NMDAR activity has been associated with elevated levels of oxidative stress[155] and Aβ production[158]. Therefore, all of the above hypotheses can feed into each other and contribute simultaneously to the observed elevated levels of DSBs in the *Sorcs2*$^{-/-}$ mice. In order to investigate this, I knocked out *SORCS2* in the LUHMES neuronal cell line and performed various experiments to address hypotheses 1 and 2. The link between SORCS2 and protection against

oxidative stress was not known until the very end of my PhD project, and thus hypothesis 3 was not investigated.

Chapter 4 Knocking out *SORCS2* in the LUHMES neuronal cell line

4.1. Introduction

To determine whether the increased levels of DNA DSBs discovered in *Sorcs2*$^{-/-}$ mice could be replicated in *SORCS2* knockout (KO) human neurons, and, if so, to address the underlying mechanism(s), I knocked out *SORCS2* in the human neuronal cell line LUHMES using CRISPR/Cas9 genome editing.

The LUHMES (Lund Human Mesencephalic) cell line is a karyotypically normal human foetal mesencephalic cell line conditionally immortalised with the v-myc oncogene[133]. Proliferation of these neuronal precursor cells can be terminated by adding tetracyclin, thus halting v-myc expression. Subsequent addition of neurotrophic growth factors results in differentiation into post-mitotic dopaminergic neurons within five days. Five to six days after initiating differentiation, the expression of stem cell and neuronal precursor markers, such as SOX2 and PAX3, is significantly reduced, while that of neuronal markers, such as synaptophysin, synapsin and PSD-95, is strongly upregulated. While the expression of most neuronal markers peaks at day 6 and remains stable thereafter, the electrophysiological properties of differentiated LUHMES cells have been shown to progressively increase with the majority of cells generating spontaneous action potentials after 10-12 days of differentiation[133]. In addition to expressing pre- and post-synaptic, and dopaminergic markers, differentiated LUHMES cells have also been shown to express proteins relevant to AD, including APP, Aβ, sAPPβ, tau, p-tau and γ-secretase[134,159].

The ability to knock out genes using CRISPR/Cas9 genome editing and examine the effect of the mutations introduced following neuronal differentiation, makes the LUHMES cell line a useful model system[135].

4.1.1. Aim

The aim of the work described in this chapter was to use CRISPR/Cas9 genome editing to generate mutant LUHMES cell lines in which the *SORCS2* gene was knocked out.

4.2. Assessing the expression of the Vps10p-domain receptor family in the LUHMES neuronal cell line

To date, no study has examined the expression of the sortilin gene family in the LUHMES neuronal cell line. In order to address this, the expression profiles of all five family members were assessed using quantitative reverse transcriptase PCR (qRT-PCR) and the TaqMan system. For each gene, a probe that amplified all (or the majority of) isoforms was selected. The expression of all genes of interest was normalised to the expression of three reference genes from a panel of eight, identified as the most stable in this sample set by the GeNorm software (Figure 4.1.). All family members, except *SORCS3*, were detected in both proliferating cells and neurons from different differentiation days (day 6, 8, 10 and 12) (Figure 4.2.). A drop in the expression levels of *SORL1*, *SORCS1* and *SORCS2* upon differentiation (day 6) was observed, while *SORT1* mRNA levels remained relatively stable up to day 10. For each of the genes detected, expression peaked at day 12 (Figure 4.2.A). Amongst the family members detected, *SORCS2* showed the highest expression in both proliferating cells and neurons, with its mRNA levels being approximately four times higher at day 12 compared to those of the rest of the family members.

To follow up the qRT-PCR results for *SORCS3*, I designed a primer pair, targeting the same *SORCS3* region as the TaqMan probe. RT-PCR with these primers detected *SORCS3* in a human brain cDNA sample, but not in cDNA samples from both proliferating and differentiated LUHMES cells (Figure 4.2.B). These results, combined with the qRT-PCR data, suggest that the *SORCS3* gene is not expressed in LUHMES cells.

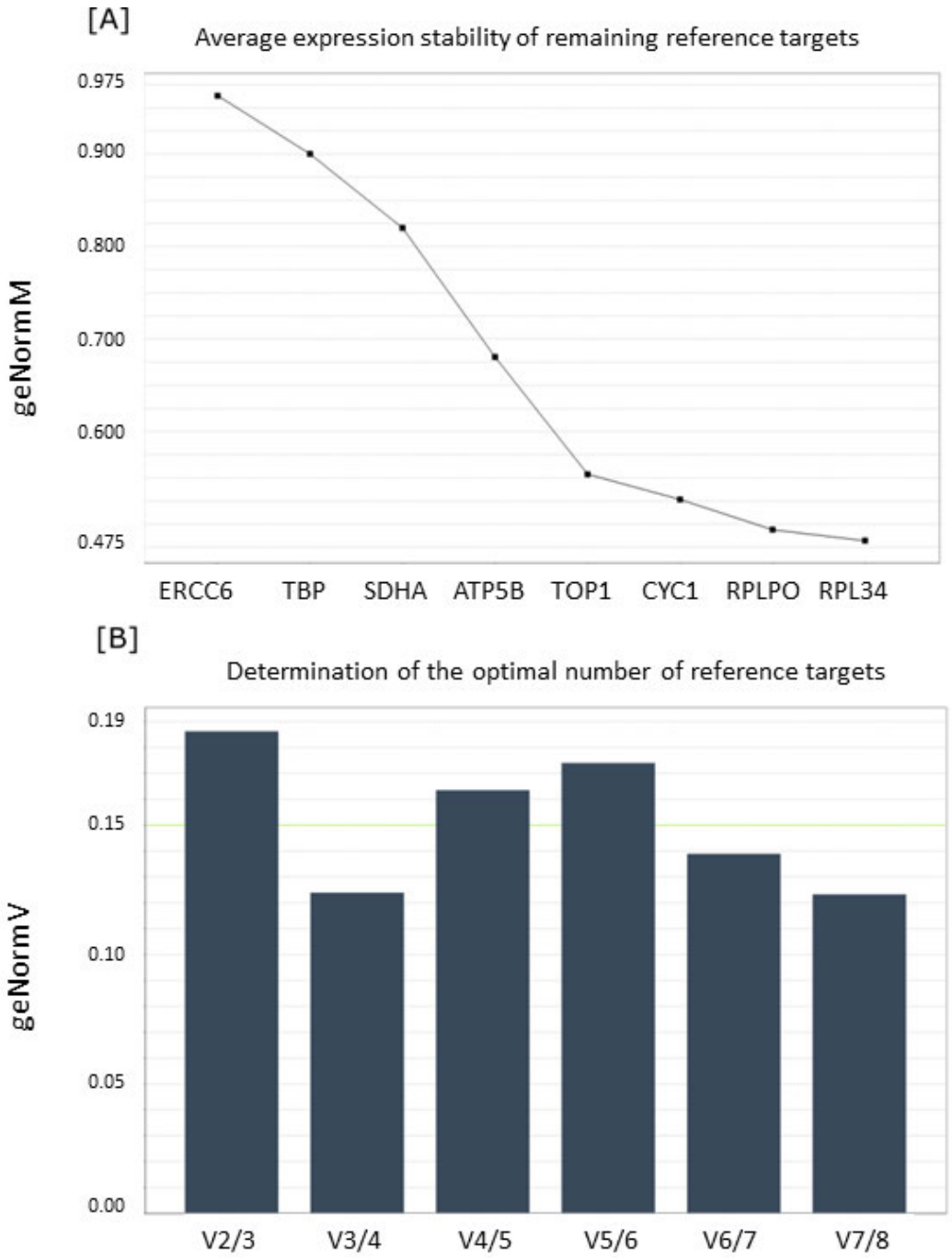

Figure 4.1. Reference gene expression stability analysed using the GeNorm software. [A] The average expression stability measure (depicted as geNorm M on the y-axis) of the eight tested reference genes (shown on the x-axis). The reference genes were ranked according to their stability from left (least stable) to right (most stable) on the x-axis. [B] Determination of the optimal number of reference genes for normalization was calculated using pairwise variation (V) analysis. V values (geNorm V) are indicated on the y-axis, and the optimal number of reference genes on the x-axis. V value of 0.15 was set as a threshold and any values lower than this indicated the number of reference genes sufficient for valid normalisation.

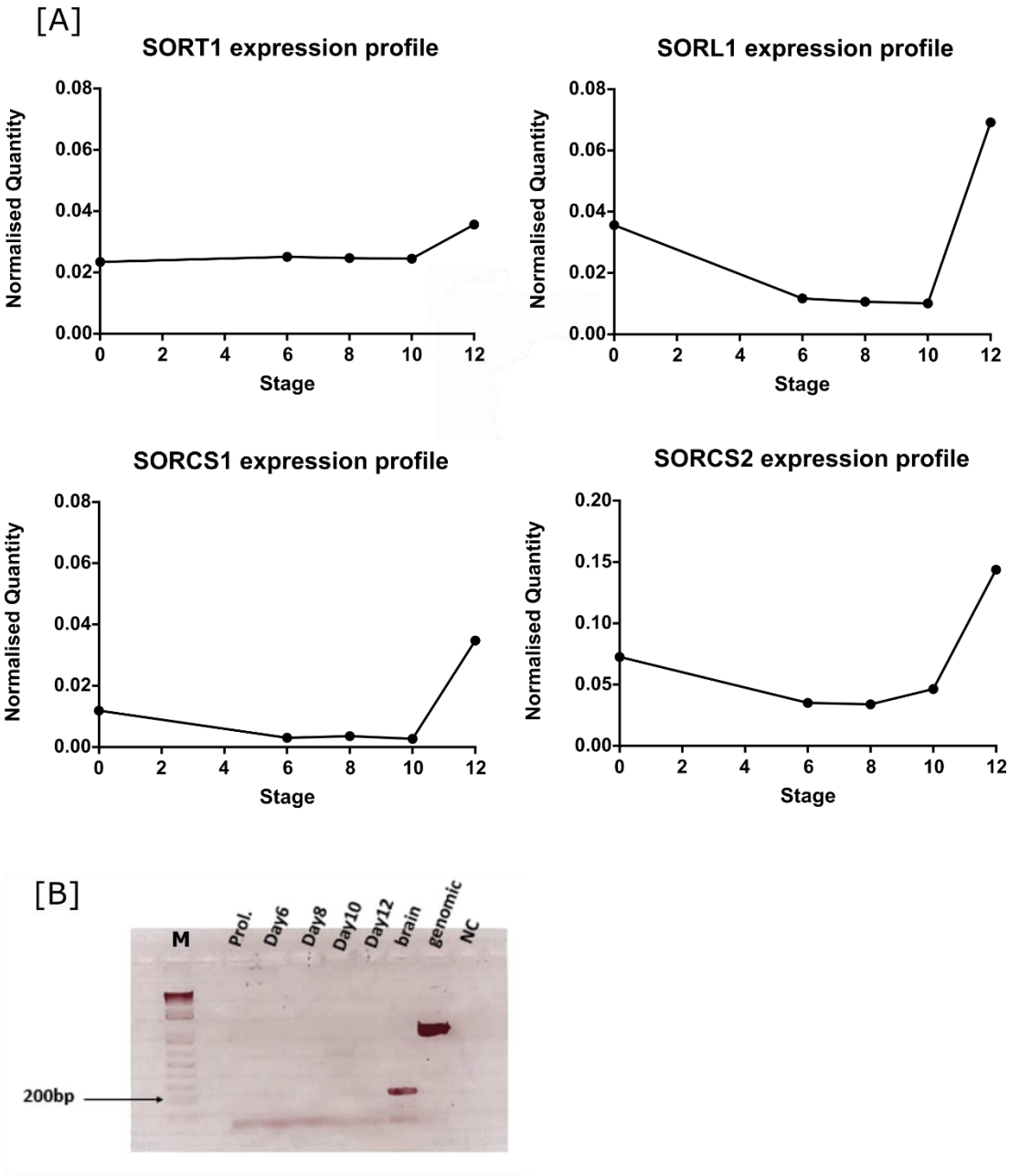

Figure 4.2. Assessing the expression of the sortilin family in proliferating and differentiated LUHMES cells. [A] SORT1, SORL1, SORCS1 and SORCS2 expression levels over the course of differentiation of proliferating LUHMES cells into neurons. The x-axis shows days of differentiation with day 0 corresponding to proliferating cells and days 6, 8, 10 and 12 to neurons at different stages of differentiation. The y-axis shows the mRNA quantity normalised to the geometric mean of the most stable reference genes selected by the GeNorm software. The experiment was performed on two independent growths, and the graphs displayed correspond to one experiment. [B] RT-PCR assessing the expression of SORCS3 in LUHMES proliferating cells and neurons. 1.5% agarose gel stained with SYBR Safe. The first track (indicated as 'M' for marker) was loaded with 1kb+ ladder, second track was left empty, tracks 3-7 were loaded with amplified cDNA from LUHMES proliferating cells and neurons of different

differentiation days. Brain cDNA (track 8) was included as positive control, and genomic DNA (track 9) as a control for genomic contamination. The last track was loaded with a non-template control (NTC). The band appearing at 200bp is the expected product size corresponding to SORCS3. No band was observed for any of the samples derived from both proliferating and differentiated LUHMES cells.

4.2.1. Detecting SORCS2 in proliferating and differentiated LUHMES cells

The qRT-PCR results indicated that *SORCS2* was expressed in both proliferating and differentiated LUHMES cells. To obtain information about protein levels, I collected protein from WT LUHMES cells at proliferating stage and two differentiation days (day 12 and day 14). Using a commercially available anti-SORCS2 antibody (AF4238, R&D), I detected a band of the expected protein size (approximately 130kDa) in both proliferating and differentiated LUHMES cells. However, in proliferating cells additional bands of different sizes were also detected. These could either constitute background bands or be suggestive of more than one *SORCS2* isoform being expressed at this cellular stage (Figure 4.3).

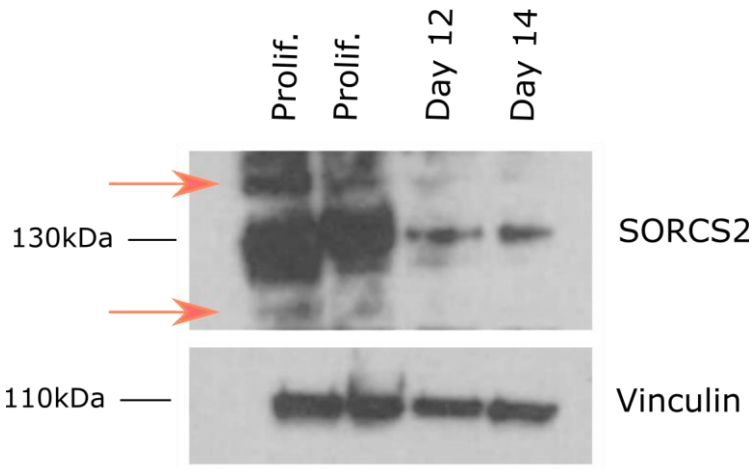

Figure 4.3. Detecting SORCS2 in proliferating and differentiated LUHMES cells. A representative western blot image, the band appearing in both proliferating and differentiated (day 12 and day 14) LUHMES cells at approximately 130kDA corresponds to SORCS2 (top panel). Red arrows point at additional bands observed in the protein samples derived from proliferating cells. Samples were probed for vinculin (approximately 110kDA, bottom panel), as well, to verify equal loading. The expected product sizes are marked on the left of the image.

4.3. Knocking out *SORCS2* in LUHMES cells using CRISPR/Cas9 genome editing

4.3.1. Experimental design

Generating functional knockouts using CRISPR/Cas9 genome editing is a multi-step process. The generic workflow is shown in Figure 4.4., and each stage is explained below.

A guide RNA (gRNA) is required to target the Cas9 protein to the genomic site of interest. There are two major considerations in the gRNA design and selection: (i) the presence of a 5'-NGG PAM site adjacent to the desired cut site and (ii) the minimisation of the gRNA's off-target activity. There are many online tools that identify suitable target sites within a genomic sequence of interest and rank the available gRNAs based on their predicted off- and on-target activity. I used two such tools: MIT Target Finder (http://crispr.mit.edu) and DNA 2.0 (https://www.dna20.com/eCommerce/cas9/input) to design gRNAs. Where possible, gRNAs ranked top by both programs were selected. When the two tools differed, gRNAs indicated by Target Finder were selected. Subsequently, I sequenced the target regions of the LUHMES genome to identify potential mismatches with the human reference genome that might compromise the gRNA activity. No mismatches were identified (data not shown).

The Cas9-expressing plasmid pSpCa9(BB)-2A-GFP was chosen as a gRNA delivery method as it carries a GFP reporter gene, allowing for the selection of successfully transfected cells using fluorescent activated cell sorting (FACS). The gRNAs selected were annealed and ligated into the plasmid (Figure 4.4.B), which was subsequently transformed into the competent Stbl3 *E.coli* strain. Ten colonies were picked for each gRNA transformation and DNA isolated. Sanger sequencing was used to confirm the correct insertion of the gRNAs in the plasmid. For each gRNA, one colony was picked, DNA was extracted and the insertion sites and gRNA sequence checked by sequencing analysis.

The pSpCa9(BB)-2A-GFP plasmid, containing the cloned gRNA, was transfected into low passage proliferating LUHMES cells using nucleofection as this has been previously reported as a preferred transfection method for human neurons [135]. The top 3% of GFP-positive cells were FAC sorted to single cells into 96-well plates (three

plates per transfection) and allowed to grow for ten days. Colonies were picked and transferred to 24-well plates and then further expanded for generating frozen stocks, isolating DNA, RNA, protein from proliferating and/or differentiated cells (Figure 4.4.C).

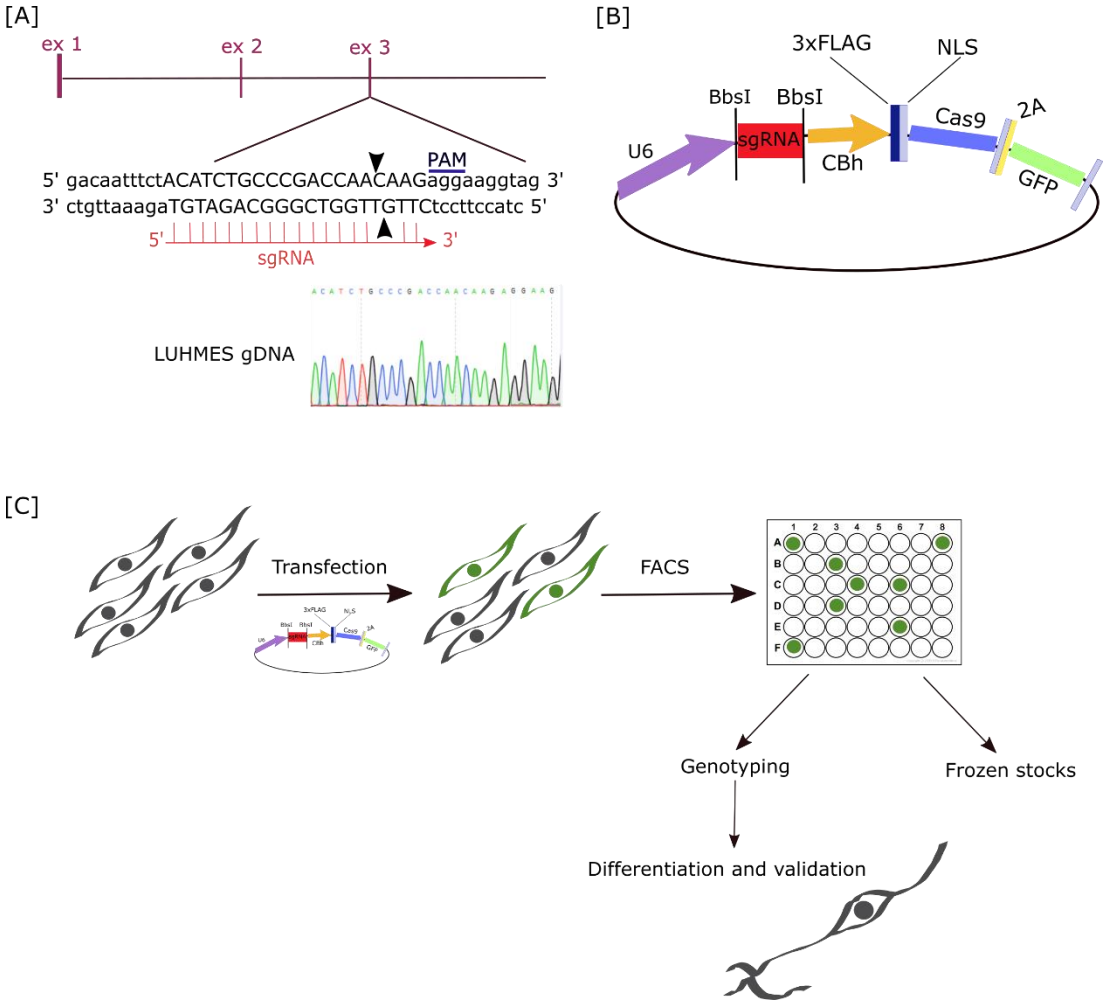

Figure 4.4. A figure showing all stages of the experimental workflow of the conducted CRISPR/Cas9 genome editing experiments: [A] gRNA selection and validation. A schematic representation of SORCS2 exon 3 selected for CRISPR/Cas9 targeting and the gRNA sequence within it. Arrowheads indicate the predicted cut site. A sequence trace showing no mismatches with the reference human genome within the gRNA sequence. [B] Schematic representation of the pSpCa9(BB)-2A-GFP plasmid, indicating the site of gRNA ligation. [C] Transfection of the pSpCa9(BB)-2A-GFP plasmid, with the ligated gRNAs, into proliferating LUHMES cells, followed by selection of the GFP-positive cells using FACS, and expansion of potentially edited clones for further validation.

4.3.1.1. Targeting the *SORCS2* gene

SORCS2 is located on chromosome 4 and has three protein-coding, one processed and five predicted transcripts (Figure 4.5). In order to achieve a complete *SORCS2* KO, I first checked which of the protein-coding and predicted transcripts were expressed in LUHMES. I designed RT-PCR primers targeting areas unique for each transcript or, when this was not possible, primers detecting a group of transcripts. No primers were designed for the processed transcript, as this type of transcripts does not contain an open reading frame and thus does not code for protein.

From all screened transcripts, the only one that was not found in either human brain or LUHMES neuron cDNA was the predicted transcript XM_005247987.4 (primer pair 4; Figure 4.6.B). All other predicted transcripts were detected in both human brain and LUHMES neuron cDNA (Figure 4.6.B). The protein-coding transcript, ENST00000329016.10, which differs from the full length *SORCS2* transcript (ENST00000507866.6) by missing the first exon, was also detected in LUHMES neurons despite its existence being poorly supported according to its TSL (transcript support level) score in the Ensembl genome browser (primer pair 8, Figure 4.6.B). Unfortunately, it was impossible to determine whether the ENST00000505529.1 transcript specifically was expressed in LUHMES as it did not have any unique regions and needed to be grouped with other *SORCS2* transcripts (primer pair 2, Figure 4.6.A). However, this transcript shows a very high TSL score according to the Ensembl genome browser. Meanwhile, ENST00000507866.6 constitutes the full-length protein-coding *SORCS2* transcript and I had previously detected SORCS2 in protein samples from differentiated LUHMES cells. Therefore, this transcript was not included in the isoform screen.

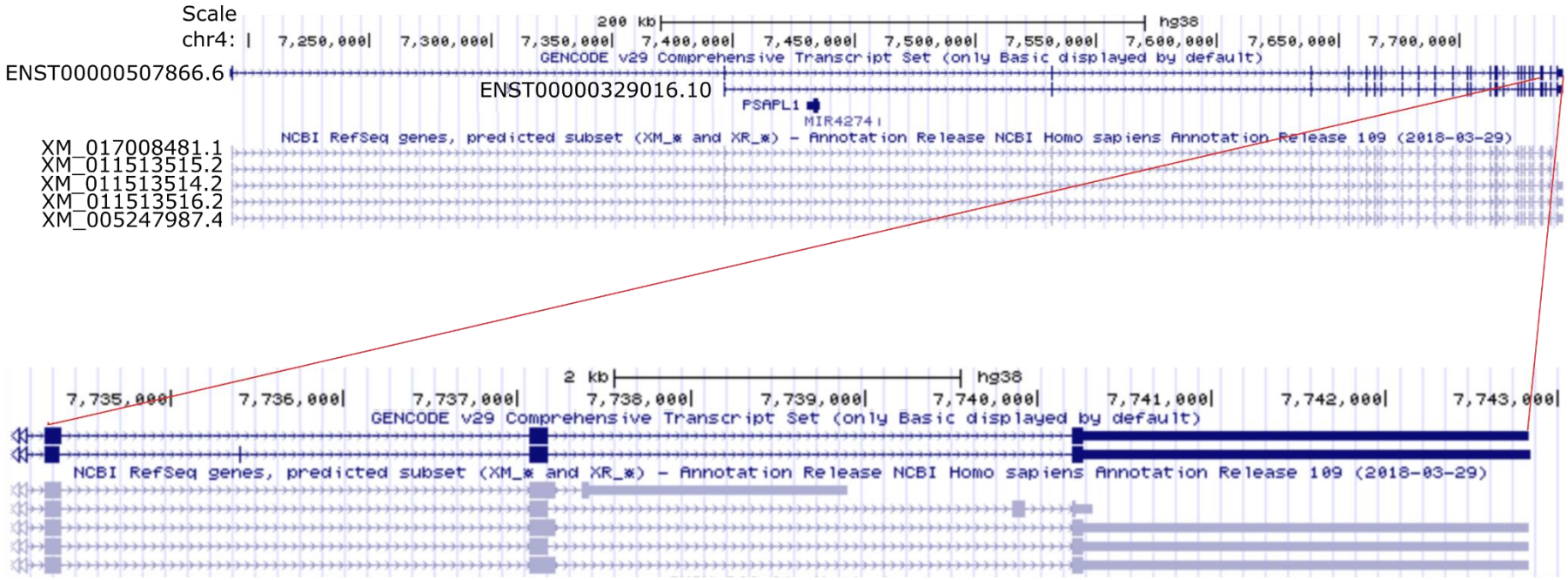

Figure 4.5. A diagram showing all protein-coding and predicted SORCS2 transcripts (apart from ENST00000505529.1, which was identified from the Ensembl genome browser). The second image shows the region where the identified SORCS2 transcripts differ from each other- exons 25-27. Individual tracks (from top to bottom) show the scale, the position in the genome, the structure of the protein-coding (dark blue) and predicted (bright blue) transcripts. The big rectangles correspond to exons, while the smaller ones- to 3'UTR ends. The thin lines with arrowheads represent intronic areas. The Names of the transcripts are listed on the left side of the top image. [Adopted after the UCSC genome browser.]

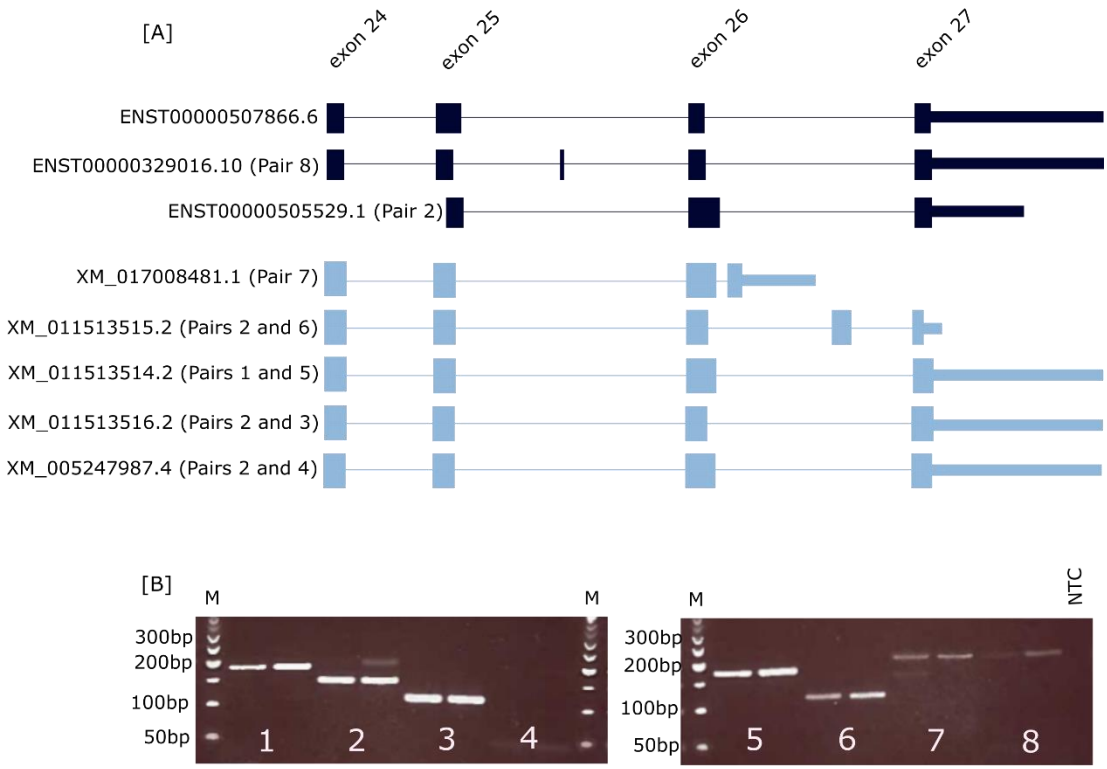

Figure 4.6. Detecting SORCS2 protein-coding and predicted transcripts in LUHMES neurons. [A] A schematic diagram of the exons where the SORCS2 transcripts differ. Protein-coding transcripts are shown in dark blue and predicted transcripts in bright blue. [B] RT-PCR with different primer pairs, designed to detect unique regions within the various SORCS2 transcripts. 1.7% agarose gel stained with SYBR Safe. 'M' indicates tracks loaded with 50bp ladder. Primer pairs are shown as 1, 2, 3, etc., and the transcript(s) amplified by a given primer pair are listed in [A]. Each pair was run on both brain cDNA (left track) and LUHMES neuron cDNA (right track). 'NTC' stands for non-template control, which included a mix of all primers used. Expected PCR product sizes: 177bp (pair 1), 145bp (pair 2), 108bp (pair 3), 184bp (pair 4), 170bp (pair 5), 114bp (pair 6), 197bp (pair 7), 192bp (pair 8).

4.3.2. Targeting *SORCS2* exon 25

I started by targeting the first exon common for all protein-coding transcripts- exon 25. I transfected proliferating cells of low passage number with the pSpCa9(BB)-2A-GFP plasmid containing gRNA, targeting this exon. The transfection efficiency was 2.17% and the top 3% of GFP-positive cells were FAC sorted into 96-well plates. From the sorted single cells, 12 clonal cell lines grew successfully and were further expanded

into 24-well plates. In the process of transferring, three of the clones died, and the rest were grown further for isolating DNA, RNA and generating frozen stocks.

4.3.2.1. Genotyping of the expanded clonal cell lines

DNA was isolated from each of the successfully expanded clones and PCR performed with primers flanking the targeted exon. PCR amplification of exon 25 suggested the presence of indels in at least four clonal lines, with one of them, clone 9, showing two bands of different sizes on the agarose gel and three clones- 1, 5 and 7, showing three bands (Figure 4.7.A). The latter three clones were discarded.

TOPO cloning, followed by Sanger sequencing, identified clone 9 as a compound heterozygote, carrying a deletion of 1bp on one allele, predicted to result in a frameshift mutation, and 6bp deletion on the other. This was the only clone predicted to carry a frameshift mutation. Sanger sequencing of the remaining five clones, showing a single band on the agarose gel (Figure 4.7.A), revealed that three of them, clones 3, 4 and 6, remained WT after the CRISPR process (hereafter referred to as 'CRISPR WT'), and one line- clone 8, was a homozygote for a 3bp deletion around the gRNA cut site. Clone 2 showed multiple sequencing traces around the predicted cut site and was discarded as a clone derived from a mixed cell population.

4.3.3.2. Assessing the expression of *SORCS2* in the CRISPR clones

In order to assess the expression of *SORCS2* in the above clones, I performed qRT-PCR on cDNA derived from proliferating cells. *SORCS2* expression levels were normalised to the expression levels of two optimum reference genes identified by the GeNorm software- *RPLP34* and *ERCC6* (data not shown). As can be seen in Figure 4.7.B, all clones showed lower *SORCS2* mRNA levels than WT LUHMES neurons grown simultaneously, including two of the three CRISRP WT lines. However, none of the other clones, including clone 9 that carried a frameshift mutation on one allele, showed mRNA lower than the CRISPR WT clones. Therefore, based on the qRT-PCR results obtained, it was concluded that this round of CRISPR targeting did not produce any *SORCS2* KO clones.

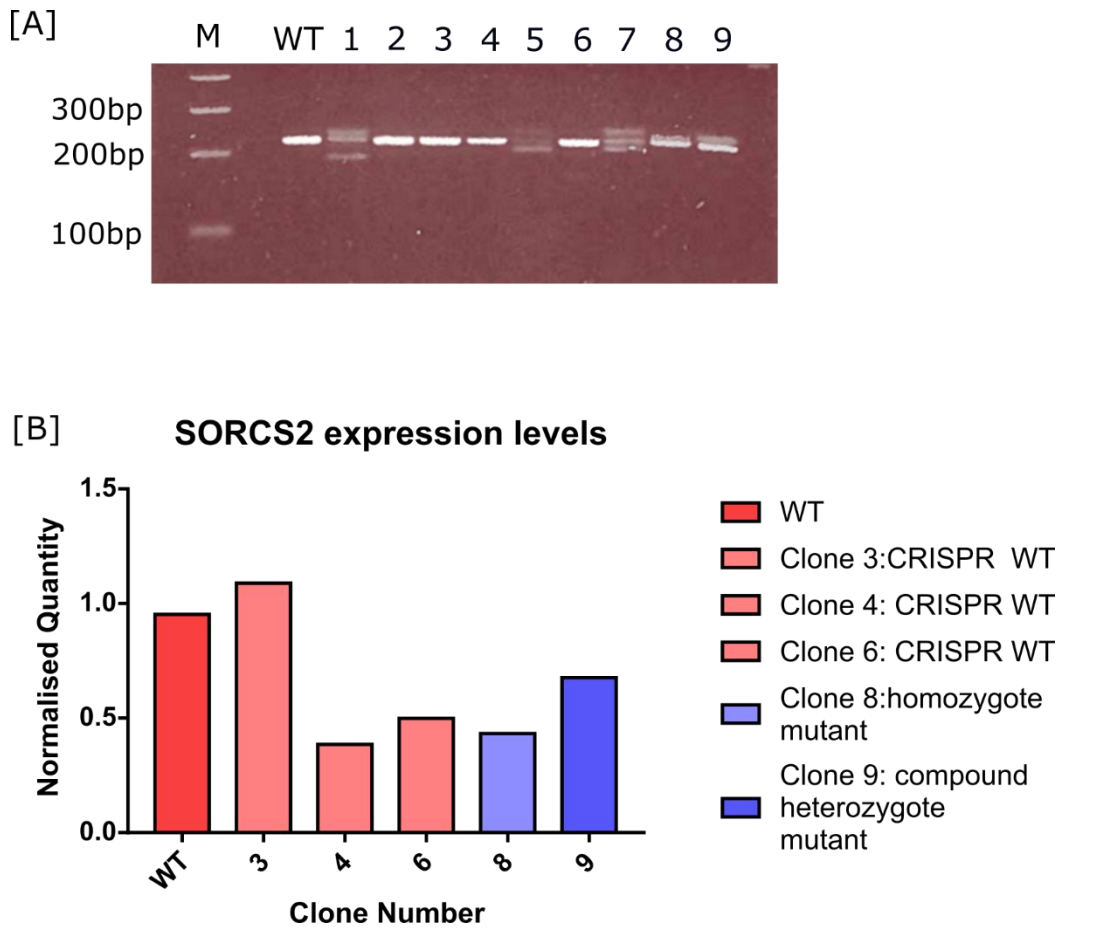

Figure 4.7. Results of CRISPR/Cas9 genome editing of SORCS2 exon 25. [A] Genotyping of the isolated CRISPR clones. PCR with primers, flanking SORCS2 exon 25, on WT genomic DNA (WT) and nine surviving clones (1-9), run on 1.7% agarose gel loaded with 1 Kb Plus ladder and stained with SYBR Safe. The first track was loaded with the ladder ('M'), the second track was left empty, the third track was loaded with PCR product from WT LUHMES cells and the rest with samples from the nine clones isolated. The expected PCR product size was 223bp. [B] qRT-PCR assessing SORCS2 mRNA levels in the isolated clones, excluding the clones that were identified as derived from mixed cell populations (clones 1, 2, 5 and 7). The x-axis shows the clones, while the y-axis shows the mRNA quantity normalised to the geometric mean of the most stable reference genes selected by the GeNorm software. The graph represents results from a single experiment.

4.3.3. Targeting *SORCS2* exon 1

As targeting *SORCS2* exon 25 did not generate *SORCS2* KO clones and because of concerns that mutations introduced in such downstream exons were less likely to lead to functional knockouts, I decided to target the first exon of the full length, protein-

coding *SORCS2* transcript ENST00000507866.6. As described previously, I transfected proliferating LUHMES cells with the pSpCa9(BB)-2A-GFP plasmid containing gRNA targeting this exon. Using this gRNA, the transfection efficiency was 7.88%. Ten days after FAC sorting of the GFP-positive cells into 96-well plates, I identified 65 clonal cell lines, which were then split into two wells of a 24-well plate. One third of the cell suspension was kept for genotyping. Three clones (1.21, 1.22 and 1.23) did not survive the transfer process.

4.3.3.1. Genotyping of the identified clonal cell lines, where *SORCS2* exon 1 was targeted

DNA was isolated from 62 clones. PCR screening with primers flanking *SORCS2* exon 1 identified 34 clonal cell lines as carrying a single band (Figure 4.8.A). Additionally, 12 clones showed two bands on the gel, and two clones- three (Figure 4.8.A). The latter were discarded. Susan Anderson performed Sanger sequencing analysis of all PCR products that showed a single band on the gel. Subsequently, I carried out the analysis of the sequencing data. High quality traces were obtained for 26 out of the 34 clones sequenced. Analysis of the traces identified one clone– 1.30, as a CRISPR WT, eight clones as heterozygotes, showing double sequence around or after the cut site, and the remaining 17 clones as homozygotes, carrying the same indel on both alleles (Table 4.1.). Subsequently, I repeated the PCR on all ten samples that did not give a product the first time, together with the ones that failed to sequence. The second round of PCR identified two clones as showing two bands on the gel, another two as carrying more than two alleles (showing three bands on the gel) and additional seven as showing a single band (Figure 4.8.B). I set up and sent those for sequencing. Two of the sequenced clones did not give clean traces. However, from the rest, three clones carried homozygous inframe deletions and one clone- a deletion within a regulatory region (Table 4.1.).

Twelve clones either failed to PCR amplify or showed unreadable sequencing traces. I therefore decided to obtain fresh DNA samples from them and repeat the genotyping. Unfortunately, the majority of clones did not recover after thawing. However, five (1.5, 1.29, 1.44, 1.51 and 1.59) clones survived. Susan Anderson isolated DNA from these clones and performed PCR amplification and Sanger sequencing. DNA isolated from clone 1.44 failed to PCR amplified even though the same amount of DNA was used as a template as for the other clones. Therefore, it

was concluded that clone 1.44 might carry a big deletion affecting the primer binding sites. The results from the genotyping are included in Table 4.1.

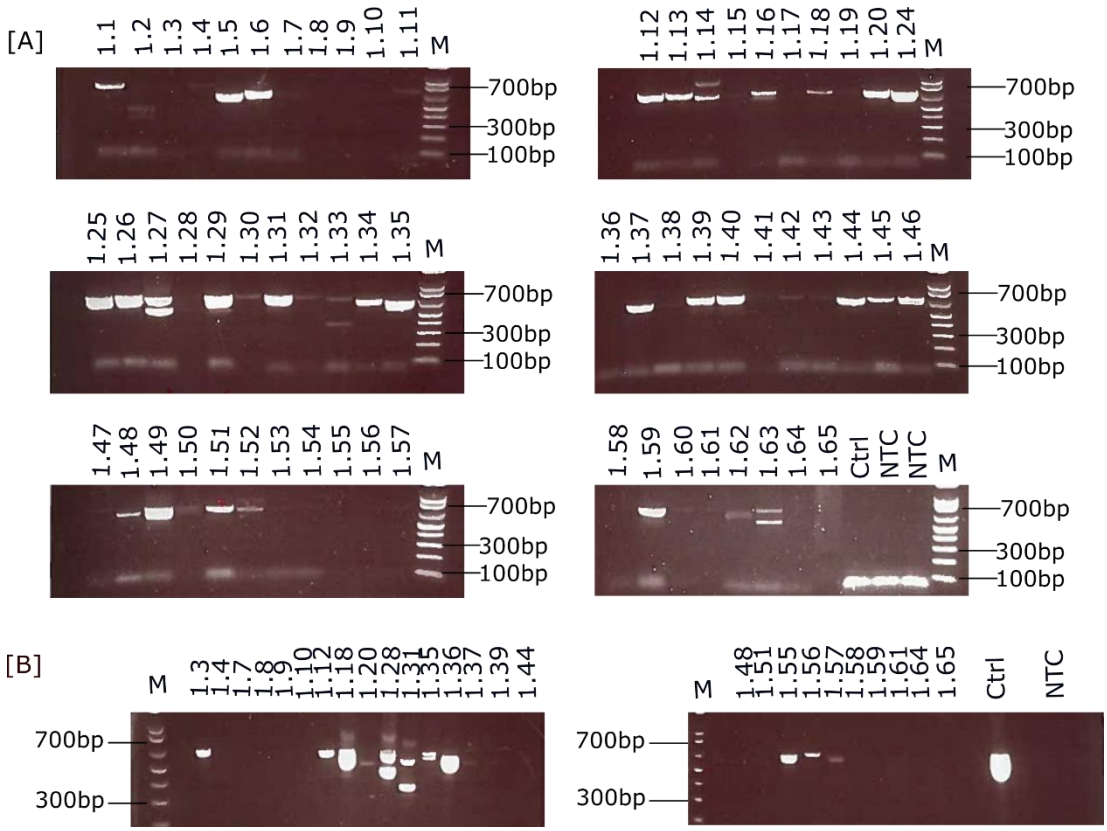

Figure 4.8. Genotyping of the 62 clones isolated after targeting SORCS2 exon 1. 1.7% agarose gels loaded with 1 Kb Plus ladder (tracks labelled with 'M') and stained with SYBR Safe. Tracks loaded with samples from the CRISPR clones are labelled as 1.1., 1.2, 1.3, etc. LUHMES genomic DNA was included as a positive control (tracks labelled as 'Ctrl'), and non-template control tracks are depicted as 'NTC'. [A] Initial PCR screening of the isolated clones with primers, flanking the targeted SORCS2 exon 1. Due to pipetting error during setting up the PCR, the positive control ('Ctrl') did not give a band of the expected product size. [B] Subsequent PCR screening on the DNA samples that failed either to amplify the first time or to give clean sequencing traces. The tracks loaded with 1 Kb Plus ladder are marked with 'M'. Expected PCR product size- 700bp.

Table 4.1. A table listing all isolated clones after targeting SORCS2 exon 1, the obtained for each PCR and sequencing results, as well as the type and the consequences of the introduced mutation.

Clone	PCR	Sequencing	Frameshift	VEP predicted impact
1.1	Single band	6bp deletion	No	Inframe deletion (moderate effect)
1.2	Three bands	Not sequenced, discarded		
1.3	Single band	6bp deletion	No	Inframe deletion (moderate effect)
1.4	Single band	Sequencing failed		
1.5	Single band	137bp deletion	Yes	Frameshift variant (high/modifier effect), affects a transcription factor binding site and a regulatory region (promoter)
1.6	Single band	85bp deletion	Yes	Frameshift variant (high/modifier effect), affects a putative regulatory region (promoter)
1.7	No PCR product			
1.8	No PCR product			
1.9	No PCR product			
1.10	No PCR product			
1.11	Single band	Double sequence trace around the cut site		
1.12	Single band	1bp mismatch + 6bp deletion	No	Inframe deletion (moderate effect)
1.13	Single band	Double sequence trace around the cut site		

Clone	PCR	Sequencing	Frameshift	VEP predicted impact
1.14	Three bands	Not sequenced, discarded		
1.15	Single band	54bp deletion	No	Inframe deletion (moderate effect)
1.16	Double band			
1.17	Single band	6bp deletion + double sequence trace around the cut site		
1.17	Single band	6bp deletion + double sequence trace around the cut site		
1.18	Single band	Sequencing failed		
1.19	Single band	1bp deletion	Yes	Frameshift mutant (high effect)
1.20	Single band	136bp deletion	No	Inframe deletion (moderate effect)
1.21	Dead			
1.22	Dead			
1.23	Dead			
1.24	Single band	Double sequence trace around the cut site		
1.25	Double band			
1.26	Double band			

Clone	PCR	Sequencing	Frameshift	VEP predicted impact
1.27	Double band			
1.28	Three bands	Not sequenced, discarded		
1.29	Single band	In progress		
1.30	Single band	WT sequence	No	WT
1.31	Three bands	Not sequenced, discarded		
1.32	Single band	Double sequence trace around the cut site		
1.33	Double band	N/A		
1.34	Double band			
1.35	Double band			
1.36	Single band	81bp deletion	No	Inframe deletion, affecting a putative promoter and a CTF-binding site (moderate/ modifier effect)
1.37	Single band	Did not work		
1.38	Single band	74bp deletion	No	Inframe deletion, affecting a putative promoter and a CTF-binding site (moderate/ modifier effect)
1.39	Single band	Did not work		
1.40	Single band	Double sequence trace around the cut site		

Clone	PCR	Sequencing	Frameshift	VEP predicted impact
1.41	Single band	Double sequence trace around the cut site		
1.42	Single band	266bp deletion	No	Inframe deletion, affecting a putative promoter and a CTF-binding site (moderate/ modifier effect)
1.43	Single band	7bp deletion	Yes	Frameshift mutation (high effect)
1.44	PCR did not work			
1.45	Single band	48bp deletion	No	Inframe deletion (moderate effect)
1.46	Double band			
1.47	Single band	Double sequence trace		
1.48	Single band	Did not work		
1.49	Double band			
1.50	Single band	Double sequence trace		
1.51	Single band	150bp deletion	Yes	Frameshift variant (high/modifier effect), affects a putative transcription factor binding site and regulatory region (promoter)
1.52	Double band			
1.53	Single band	26bp deletion	Yes	Frameshift mutation (high effect)
1.54	Single band	9bp deletion	No	Inframe deletion, affecting a putative promoter and a CTF-binding site (moderate/modifier effect)
1.55	Dead			
1.56	Single band	6bp deletion	No	Inframe deletion (moderate effect)

Clone	PCR	Sequencing	Frameshift	VEP predicted impact
1.57	Single band	Did not work		
1.58	Did not work			
1.59	Single band	2bp insertion	Yes	Frameshift mutation (high effect)
1.60	Single band	12bp deletion	No	Inframe deletion
1.61	Single band	Did not work		
1.62	Single band	Did not work		
1.63	Double band			
1.65	PCR did not work			

4.3.3.2. Identifying *SORCS2* knockouts amongst the clones, carrying variants predicted to have an effect on the protein sequence

Given the fact that mRNA levels do not necessarily correspond to protein levels[160], I decided to screen for potential *SORCS2* KOs using western blot analysis. Normally, inframe deletions do not affect translation, while frameshift mutations alter the reading frame and ultimately introduce a premature stop codon. Bearing this in mind, I screened all clones that carried a frameshift mutation (1.5, 1.6, 1.19, 1.29, 1.44, 1.51, 1.53 and 1.59), as well as clones 1.42 and 1.44. Clone 1.42 had a large inframe deletion that was predicted to affect a putative promoter and a CTF-binding site by the Ensembl VPN tool. Meanwhile, DNA isolated from clone 1.44 failed PCR amplify and was thus assumed to carry a big deletion. Clone 1.30, which remained a WT after the CRISPR process (CRISPR WT), was also included in the analysis as a control.

The selected CRISPR-edited clones were differentiated into neurons alongside WT LUHMES cells of a similar passage number. During the process of differentiation, clone 1.6 and 1.19 died and I was unable to revive the remaining frozen vials. After 14 days of differentiation, I collected protein samples from all of the remaining clones, which I then used to analyse SORCS2 expression using western blot. Detecting SORCS2 with a commercially available antibody (AF4238, R&D) identified clones 1.5, 1.29, 1.44, 1.51, 1.59 (Figure 4.9.B) and 1.42 and 1.53 (Figure 4.9.A) as *SORCS2* KOs. Meanwhile, the CRISPR WT clone 1.30 showed protein levels close to that of the control WT LUHMES cell line as determined by visual inspection (Figure 4.9.A). These results were confirmed using a second, custom-made antibody (F7100, kind gift of Assoc. Prof Simon Glerup), specific for SORCS2 (Figure 4.9.A). However, when probed with this antibody, clone 1.42 showed a faint band of the correct size suggesting that it is not a knockout.

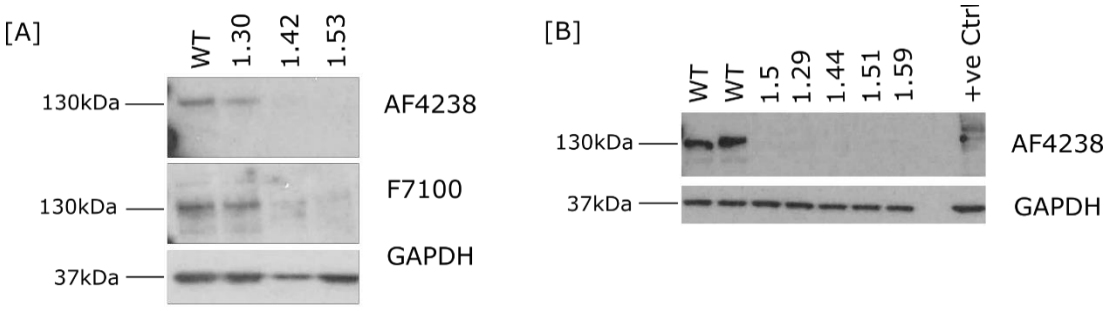

Figure 4.9. Western blot analysis identifying SORCS2 KO clones amongst the ones predicted to carry mutations, affecting the protein sequence. [A] Representative western blot images, indicating clone 1.53, but not clone 1.42 as SORCS2 KO. Protein samples from WT LUHMES

and the CRISPR WT clone 1.30 were included as controls. Clone 1.30 showed SORCS2 protein levels similar to the WT. Probing the same membrane with two different antibodies against SORCS2- AF4238 (top panel) and F7100 (middle panel) confirmed the results. Samples were probed for GAPDH (bottom panel), as well, to verify equal loading. The expected product sizes (SORCS2- 130kDa and GAPDH- 37kDA) are marked on the left of the images. [B] Representative image of a western blot, indicating clones 1.5, 1.29, 1.44, 1.51 and 1.59 as SORCS2 KOs. Two LUHMES WT lines of a similar passage number and grown at the same time were included as controls. Protein lysate from HEK cells, carrying a SORCS2 expression plasmid, was used a positive control. Probing for GAPDH verified equal loading. The expected product sizes (SORCS2- 130kDa and GAPDH- 37kDA) are marked on the left of the images. Western blots were repeated on at least three independently collected protein lysates from each clone.

4.3.4. Targeting *SORCS2* exon 3

I decided to target exon 3 by CRISPR/Cas9 genome editing simultaneously, as one *SORCS2* transcript, ENST00000329016.10, which lacks exon 1 was also detected in LUHMES neurons using RT-PCR (Figure 4.6.). As exon 2 was too short for gRNA design, I targeted *SORCS2* exon 3. I transfected proliferating LUHMES cells with gRNA homologous to a region within this exon. The efficiency of the transfection was considerably lower- 1.33%, compared to that obtained when targeting exon 1 (7.88%). However, after ten days, I identified 73 clonal cell lines. Each line was split into two wells of a 24-well plate and one third of the cell suspension was kept for genotyping.

4.3.4.1. Genotyping of the identified clonal cell lines where *SORCS2* exon 3 was targeted

In order to genotype the 73 clones where exon 3 was targeted, Susan Anderson set up a PCR assay using intronic primers flanking the targeted exon. Initial screening identified 48 clonal cell lines as potentially homozygous, showing a single band on the gel. Fifteen clones had two bands, and five others were discarded due to the presence of three bands (Figure 4.10.).

All 57 samples where there was a single band on the gel were subjected to sequencing analysis by Susan Anderson, and I analysed the results. Of the 57 sequenced samples, 18 had double traces in the region of the predicted gRNA cut site (Table 4.2). The VEP tool was used as before to predict the effect of the mutations on the protein sequence of the remaining 39 clones. Twelve clones carried inframe

deletions, ten clones had frameshift mutations and three clones carried a mutation affecting a splice donor site, as well as an enhancer region. The latter were predicted to have a damaging effect by VEP. Another clone was heterozygous for one base pair mismatch (Table 4.2). Additionally, the DNA extraction failed for one clone, and the PCR assay did not work for four others (Figure 4.10). I decided not to repeat the genotyping of the failed samples since I had identified enough clones that carried frameshift mutations.

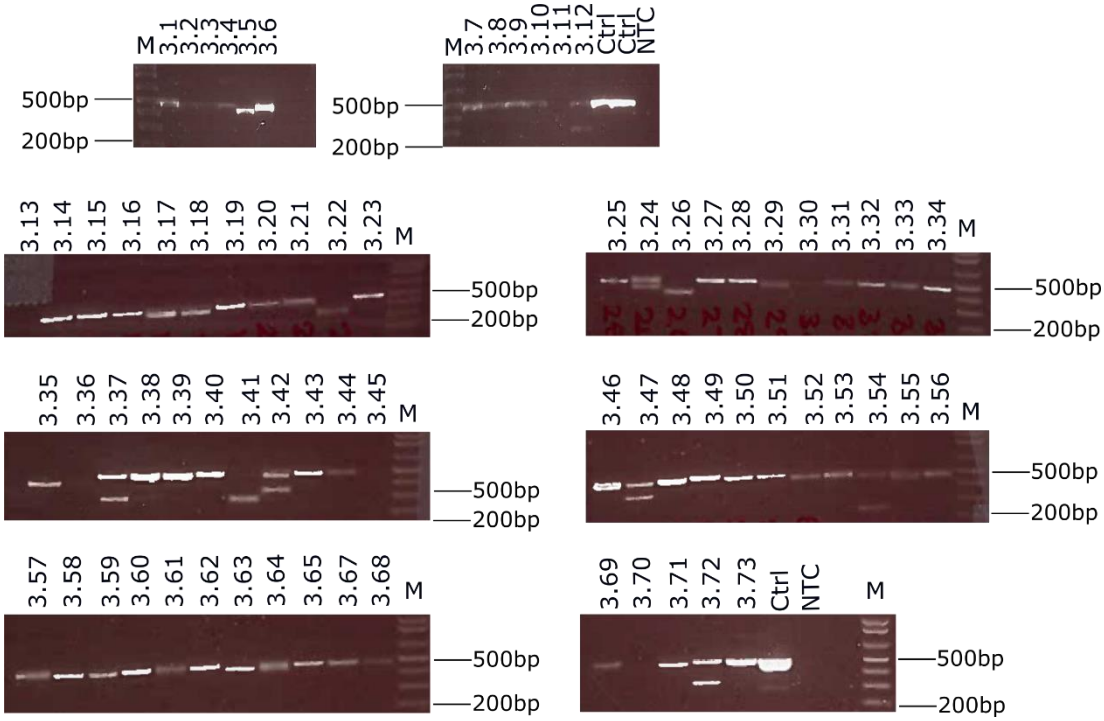

Figure 4.10. Genotyping of the 73 clones isolated after targeting SORCS2 exon 3, using primers, flanking the targeted exon. 1.7% agarose gel stained with SYBR Safe. Tracks labelled with 'M' were loaded with 1 Kb Plus ladder. Tracks loaded with samples from the CRISPR clones are labelled from 3.1. to 3.73. LUHMES genomic DNA was included as a positive control (tracks labelled as 'Ctrl'). 'NTC' corresponds to non-template control. The expected PCR product size was 476bp.

Table 4.2. A table listing all isolated clones after targeting SORCS2 exon 3, the obtained for each PCR and sequencing results, as well as the type and the consequences of the introduced mutation.

Clone	PCR	Sequencing	Frameshift	VEP predicted impact
3.1	Single band	Double sequence trace around the cut site		
3.2	Single band	9bp deletion	No	Inframe deletion (moderate effect)
3.3	Single band	Double sequence trace around the cut site		
3.4	Single band	1bp insertion	Yes	Frameshift mutation (high impact)
3.5	Single band	33bp deletion	No	Inframe deletion (moderate effect)
3.6	Single band	Double sequence trace around the cut site		
3.7	Single band	Double sequence trace around the cut site		
3.8	Single band	Double sequence trace around the cut site		
3.9	Single band	Mismatches around the cut site		
3.10	Single band	Double sequence trace around the cut site		
3.11	PCR did not work			
3.12	Double band			
3.13	PCR Did not work			
3.14	Single band	Double sequence trace around the cut site		
3.15	Single band	Double sequence trace around the cut site		
3.16	Single band	12bp deletion	No	Inframe deletion (moderate effect)

Clone	PCR	Sequencing	Frameshift	VEP predicted impact
3.17	Double band			
3.18	Double band			
3.19	Single band	Double sequence trace around the cut site		
3.20	Single band	12bp deletion	No	Inframe deletion (moderate effect)
3.21	Single band	9bp deletion	No	Inframe deletion (moderate effect)
3.22	Single band	126bp deletion	No	Donor splice site variant (high impact)
3.23	Single band	Double sequence trace around the cut site		
3.24	Three bands	Not sequenced, discarded		
3.25	Single band	Double sequence trace around the cut site		
3.26	Single band	95bp deletion	Yes	Donor splice site variant (high impact)
3.27	Single band	Double sequence trace around the cut site		
3.28	Single band	Double sequence trace		
3.29	Single band	Double sequence trace around the cut site		
3.30	Single band	Sequencing did not work		
3.31	Single band	1bp insertion	Yes	Frameshift mutation (high impact)
3.32	Single band	3bp deletion	No	Inframe deletion (moderate effect)
3.33	Single band	3bp deletion	No	Inframe deletion (moderate effect)
3.34	Single band	9bp deletion	No	Inframe deletion (moderate effect)

Clone	PCR	Sequencing	Frameshift	VEP predicted impact
3.35	Single band	56bp deletion	No	Inframe deletion (moderate effect)
3.36	Single band	3bp deletion	Yes	Frameshift mutation (high impact)
3.37	Double band			
3.38	Single band	1bp mismatch		
3.39	Single band	10bp deletion	Yes	Frameshift mutation (high impact)
3.40	Single band	2bp insertion	Yes	Frameshift mutation (high impact)
3.41	Single band	191bp deletion	Yes	Donor splice site variant (high impact)
3.42	Double band			
3.43	Single band	Double sequence trace around the cut site		
3.44	Single band	1bp insertion	Yes	Frameshift mutation (high impact)
3.45	Single band	Sequencing did not work		
3.46	Three bands	Not sequenced, discarded		
3.47	Double band			
3.48	Single band	Double sequence trace		
3.49	Single band	Multiple insertions		
3.50	Single band	6bp deletion	No	Inframe deletion (moderate effect)
3.51	Single band	1bp insertion	Yes	Frameshift mutation (high effect)
3.52	Single band	13bp deletion	Yes	Frameshift mutation (high effect)
3.53	Single band	Double sequence trace around the cut site		

Clone	PCR	Sequencing	Frameshift	VEP predicted impact
3.54	Double band			
3.55	Three bands	Not sequenced, discarded		
3.56	Single band	Double sequence trace around the cut site		
3.57	Single band	Double sequence trace around the cut site		
3.58	Single band	Double sequence trace around the cut site		
3.59	Single band	Double sequence trace around the cut site		
3.60	Single band	1bp insertion	Yes	Frameshift mutation (high impact)
3.61	Single band	Double sequence trace around the cut site	N/A	N/A
3.62	Single band	Double sequence trace around the cut site		
3.63	Single band	17bp deletion	Yes	Frameshift mutation (high impact)
3.64	Three band	Not sequenced, discarded		
3.65	Single band	Double sequence trace around the cut site		
3.66	Single band	6bp deletion	No	Inframe deletion (moderate effect)
3.67	Failed DNA isolation			
3.68	Single band	Double sequence trace around the cut site		
3.69	Single band	Double sequence trace around the cut site		
3.70	Did not work			
3.71	Single band	12bp deletion	No	Inframe deletion (moderate effect)
3.72	Double band			
3.73	Single band	Double sequence trace around the cut site		

4.3.4.2. Identifying *SORCS2* knockouts amongst the clones carrying variants predicted to have an effect on the protein sequence

As mentioned above, lines carrying a frameshift mutation are mostly likely to be knockouts. Therefore, I grew up and differentiated the following clones into neurons: 3.36, 3.39, 3.40, 3.52, 3.60 and 3.63, which carried frameshift mutations and clones 3.22, 3.26 and 3.41, which had mutations predicted to exert damaging effect on translation. LUHMES WT neurons of a similar passage number were cultured in parallel to the edited lines. During the process of differentiation, clones 3.22, 3.52 and 3.60 died and I was unable to revive the remaining from them frozen vials.

After 14 days of differentiation, I collected protein samples from all clones, which I then probed for SORCS2 using western blot. Detecting SORCS2 with the commercially available antibody AF4238 (R&D) identified clones 3.26, 3.36, 3.39, 3.40, 3.41 and 3.63 as knockouts (Figure 4.11.).

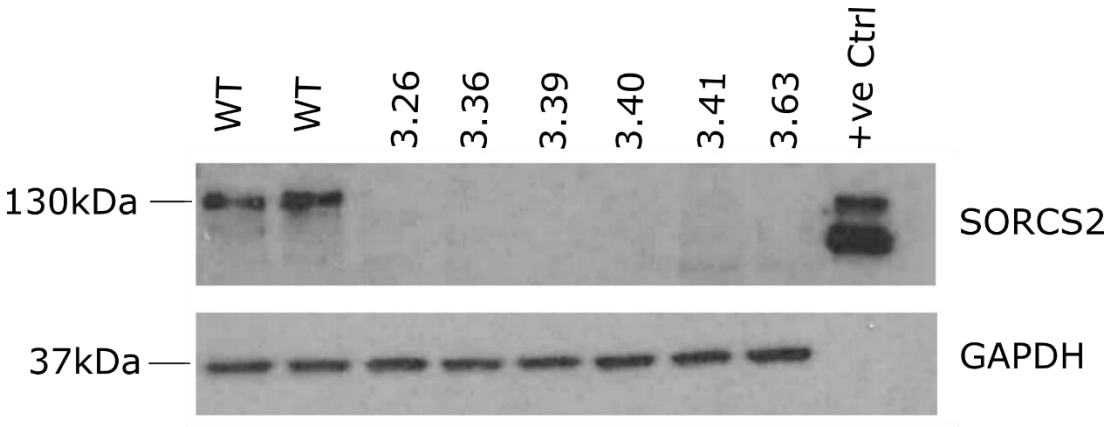

Figure 4.11. Western blot analysis identifying SORCS2 KO clones. A representative western blot image (top panel), showing clones 3.26, 3.36, 3.39, 3.40, 3.41 and 3.63 as SORCS2 KOs. Protein samples from WT LUHMES neurons of a similar passage number were included as controls. Samples were probed for GAPDH (bottom panel) to verifying equal loading. Protein lysate from HEK cells carrying a SORCS2 expression plasmid was used a positive control. The expected protein sizes (SORCS2- 130kDa and GAPDH- 37kDA) are marked on the left of the images. Western blots were repeated on at least three independently collected protein lysates from each clone.

4.4. Discussion

The aim of the work described in this chapter was to generate knockouts of the *SORCS2* gene in LUHMES cells using the CRISPR/Cas9 genome editing system.

Using qRT-PCR, I have shown that LUHMES neurons express four members of the sortilin family in what appears to be developmentally regulated manner. Additionally, RT-PCR detected potentially all, including predicted, *SORCS2* protein coding transcripts in LUHMES neurons, apart from XM_005247987.4. However, since some transcripts did not have unique regions, I needed to group them together with others. Thus, it was impossible to predict whether the ENST00000505529.1 transcript specifically was expressed in LUHMES.

After establishing which *SORCS2* isoforms were likely to be expressed in LUHMES neurons, I designed gRNAs aiming to knock out the gene. I started with a gRNA that targeted the first common exon for all transcripts (including ENST00000505529.1)-exon 25. Only 12 single cell colonies survived following sorting, and three out of these 12 clones remained WT after the CRISPR process. Meanwhile, only one (clone 9) carried a frameshift mutation only on one allele of the *SORCS2* gene. The mRNA levels detected in this clone were not lower than those detected in the clones that remained WT after the CRISPR process. Moreover, two out of the three CRISPR WT clones showed reduced *SORCS2* mRNA levels when compared to a WT cell line. This unexpected finding could be explained by either gRNA off-target activity or intrinsic cell line variability.

There are a few possible reasons for the unsuccessful attempt of knocking out *SORCS2* by targeting exon 25. First, the low number of single cell colonies identified, combined with the low proportion of clones carrying a mutation in the targeted area suggests that the used gRNA was either toxic and/or had lower efficiency. As mentioned in the introduction, two online tools were used for gRNA design- DNA 2.0 and the MIT Target Finder. When these tools ranked different gRNAs as best, the ones indicated by Target Finder were selected. This was the case when designing the gRNA targeting *SORCS2* exon 25. However, following this experiment, problems in the Target Finder algorithm were announced. Since then, many online tools predicting the efficiency of the available gRNA for a given area have been designed. Using one such tool (http://chopchop.cbu.uib.no/index.php), I found out that all available gRNA for *SORCS2* exon 25, including the one I used, display low efficiency and are predicted to have many off-targets.

Even if the gRNA used was efficient, it would have been still difficult to achieve a complete *SORCS2* KO by targeting exon 25. Frameshift mutations close to the carboxy terminal are known to be less deleterious compared to these introduced next to a translation start site (TSS), as they are more likely to produce truncated proteins rather than to activate NMD pathways[161]. Thus, in order to achieve functional *SORCS2* KO, I designed gRNAs targeting exon 1 and 3. Exons 1 and 3 are located in the beginning of all protein coding *SORCS2* transcripts (including the predicted ones), apart from ENST00000511199.1 which lacks exon 1 and ENST00000505529.1 which starts from exon 25. However, it is unlikely that these transcripts result in protein products as, according to Ensembl, they have very high TSL scores (TSL5 and TSL4, respectively) and thus are not very well supported by mRNA and/or expressed sequence tags (ESTs).

Targeting *SORCS2* exon 1 and 3 resulted in a large number of clonal lines derived from single cells, most of which were edited (only one clonal line remained CRISPR WT after targeting *SORCS2* exon 1 and none after targeting exon 3). From these, I identified a number of *SORCS2* KO clonal lines, derived from each of the used gRNAs: clones 1.5, 1.29, 1.44, 1.51, 1.53 and 1.55 (derived from gRNA targeting *SORCS2* exon 1), and 3.26, 3.36, 3.39, 3.40, 3.41 and 3.63 (derived from gRNA targeting *SORCS2* exon 3). It is highly likely that these clones constitute complete knockouts, as no bands of the expected protein size of 130kDA were observed when using two different antibodies against the human SORCS2 protein. Moreover, since targeting exon 1 knocked out *SORCS2*, it is highly unlikely that the ENST00000511199.1 transcript is translated in LUHMES. This is in agreement with the high TSL assigned to it by Ensembl.

The majority of the identified single cell colonies constituted compound heterozygotes as they showed more than one allele after PCR amplification. Thus, it is possible that some of these compound heterozygotes are knockouts for the *SORCS2* gene, as well. However, since I had already identified a reasonable number of *SORCS2* KOs and also due to time restraints, this was not further investigated. Additionally, some clones showed three bands on the PCR gel. These were discarded as they were most probably derived from more than one single cell, thus representing mixed cell populations.

Chapter 5 Assessing the effect of knocking out *SORCS2* on DNA DSB formation and potential underlying mechanisms *in vitro*

5.1. Introduction

As discussed in chapter 3, I detected increased levels of DNA DSB formation in the dentate gyrus of *Sorcs2*$^{-/-}$ mice compared to WT mice. Following on from this result, I first wanted to see whether these findings can be replicated both in the human neuronal cell line LUHMES, as well as in mouse hippocampal primary neurons. Additionally, I wanted to explore the effect of knocking out *SORCS2 in vitro* on the generation of DNA DSBs following treatment with etoposide. Etoposide is a chemotherapeutic agent that binds to DNA topoisomerase II (TopoII), trapping the enzyme in a covalent complex with the cleaved DNA and preventing break re-ligation[162]. Thus, its application allows the accumulation of DNA damage. There are two distinct isoforms of TopoII expressed in the mammalian system- TopoIIα and TopoIIβ. The former is mainly expressed in dividing cells, whereas the latter is expressed in postmitotic cells, such as neurons, where it is primarily involved in transcription-related processes[163]. Recently, TopoIIβ has also been implicated in the formation of DNA DSBs following neuronal activity and has been shown to be required for the expression of early-response genes[118]. Additionally, treatment of primary neurons with etoposide altered the expression of nearly 700 genes, with the majority of them being downregulated. However, twelve genes were upregulated, including the early-response genes *Fos*, *FosB* and *Npas4*[118]. Secondly, if knocking out *SORCS2* was found to be associated with increased DSB formation, I wanted to explore potential mechanisms underlying this finding. Suberbielle *et al.* (2013)[107] showed that treatment with Aβ oligos led to increased DSBs in WT primary neurons, and that this process was dependent on NR2B-containing NMDAR activity. Meanwhile, NR2A-containin NMDARs were required for DSB repair[107]. Similarly to etoposide, stimulating primary neuronal cultures with the synthetic NMDAR agonist, NMDA, was associated with the formation of DNA DSBs within early-response genes, including *Fos*, *FosB*, *Npas4*, and *Egr1*, as well as within genes encoding for other transcription factors and a few non-coding RNAs[118]. In mice, SorCS2 is required for the induction of hippocampal NMDAR-dependent LTP and LTD[5], as well as for the

trafficking of the NMDA receptor subunit NR2A in striatal neurons[51]. Furthermore, knocking down *SORCS2* in HEK293 cells has been previously shown to upregulate APP processing[46]. Thus, both increased Aβ production and/or dysregulated NMDA receptor signalling would constitute potential mechanisms underlying any increase in DNA DSB formation observed in the knockout cell lines. Additionally, elevated extracellular dopamine levels have been linked to increased DNA damage and neurodegeneration *in vitro*[164,165]. Therefore, given the role of SorCS2 in the development of the dopaminergic system[6], the levels of extracellular dopamine secreted by the *SORCS2* KO LUHMES neurons were also investigated as a potential source of DNA damage.

5.1.1. Aims

The aims of the experiments described in this chapter were to:

- Investigate the levels of DNA DSB formation in the generated *SORCS2* KO LUHMES lines both with and without treatment with etoposide.
- Investigate the levels of DNA DSB formation in untreated and etoposide-treated primary hippocampal neurons derived from *Sorcs2*$^{-/-}$ and WT pups.
- Investigate potential mechanisms that could account for the observed increase in DNA DSB formation upon knocking out *SORCS2 in vivo* and *in vitro*.

5.2. Assessing the levels of DNA DSB formation in the *SORCS2* KO LUHMES lines

5.2.1. Assessing DNA DSB formation in the *SORCS2* KO LUHMES lines

In order to assess whether the higher levels of DSBs observed in the *Sorcs2*$^{-/-}$ compared to WT mice could be replicated in human neurons, I grew and differentiated *SORCS2* KO (n=4) and WT (n=3) LUHMES lines of a similar passage number. Clones where *SORCS2* exon 3 was disrupted were chosen for this initial experiment, as this meant that both the long and shorter isoform of *SORCS2* were targeted. Following 14 days of differentiation, neurons were fixed and stained for γH2A.X and 53BP1. As

discussed in chapter 1, γH2A.X staining is considered the 'gold standard' for detecting DNA DSBs[111]. Meanwhile, 53BP1 plays a critical role in the cell response to DNA damage, being quickly recruited to DSB sites, where it binds to γH2A.X and acts as a scaffold for the binding of additional DNA repair proteins[166,167]. Thus, co-labelling of γH2A.X-positive foci with 53BP1 provides clear evidence for the presence of DNA DSBs.

As previously reported for primary cortical cultures, untreated LUHMES neurons showed very few γH2A.X/53BP1-positive foci, with 53BP1 appearing mostly as defuse staining (Figure 5.1.A). Approximately 100 nuclei (from four images belonging to different regions of the same coverslip) were counted for each WT and *SORCS2* KO line and the number of γH2A.X/53BP1-positive foci per nucleus was calculated. Nuclei showing apoptotic features (i.e. condensed and 'blebbing' nuclei), as well as those completely covered with intense, bright γH2A.X signal were excluded from the analysis (Figure 5.1.A). More than 90% of the analysed untreated WT neurons had fewer than 3 foci per nucleus with the majority of them showing 0 foci. These observations were consistent with previous studies of both untreated neurons[115] and other non-neuronal cell types[122]. While the *SORCS2* KO lines (n=4) had on average a greater number of γH2A.X/53BP1-positive foci per nucleus compared to WT controls (n=3), the difference was not significant (p=0.085; Figure 5.1.B). This result might have reflected a true lack of difference between the two groups or, alternatively, the small sample size might have led to insufficient power to detect a difference.

To overcome these limitations, I repeated the experiment by including additional WT and *SORCS2* KO clones in the analysis. This time, *SORCS2* KO clones generated after targeting exon 1 (n=4) were added to the ones where exon 3 was disrupted (n=5) in an attempt to ensure that the observed phenotype is not the result of CRISPR off-target activity, as any off-targets will differ between two gRNAs. As before, WT (n=9) and *SORCS2* KO (n=9) clones were differentiated until day 14, fixed and stained for γH2A.X and 53BP1. Approximately 100 nuclei (from four images taken from the same coverslip) were examined for each cell line, and the number of foci where γH2A.X and 53BP1 co-localised was counted. As for the initial experiment, there was no significant difference in the number of foci per nucleus between the WT and the *SORCS2* KO clones (p=0.647, Figure 5.1.C). No significant difference was observed also when the number of γH2A.X/53BP1-positive foci per nucleus was compared between the

SORCS2 KO clone generated after targeting exon 1 and those obtained after targeting exon 3 (p= 0.7780, Figure 5.1.D).

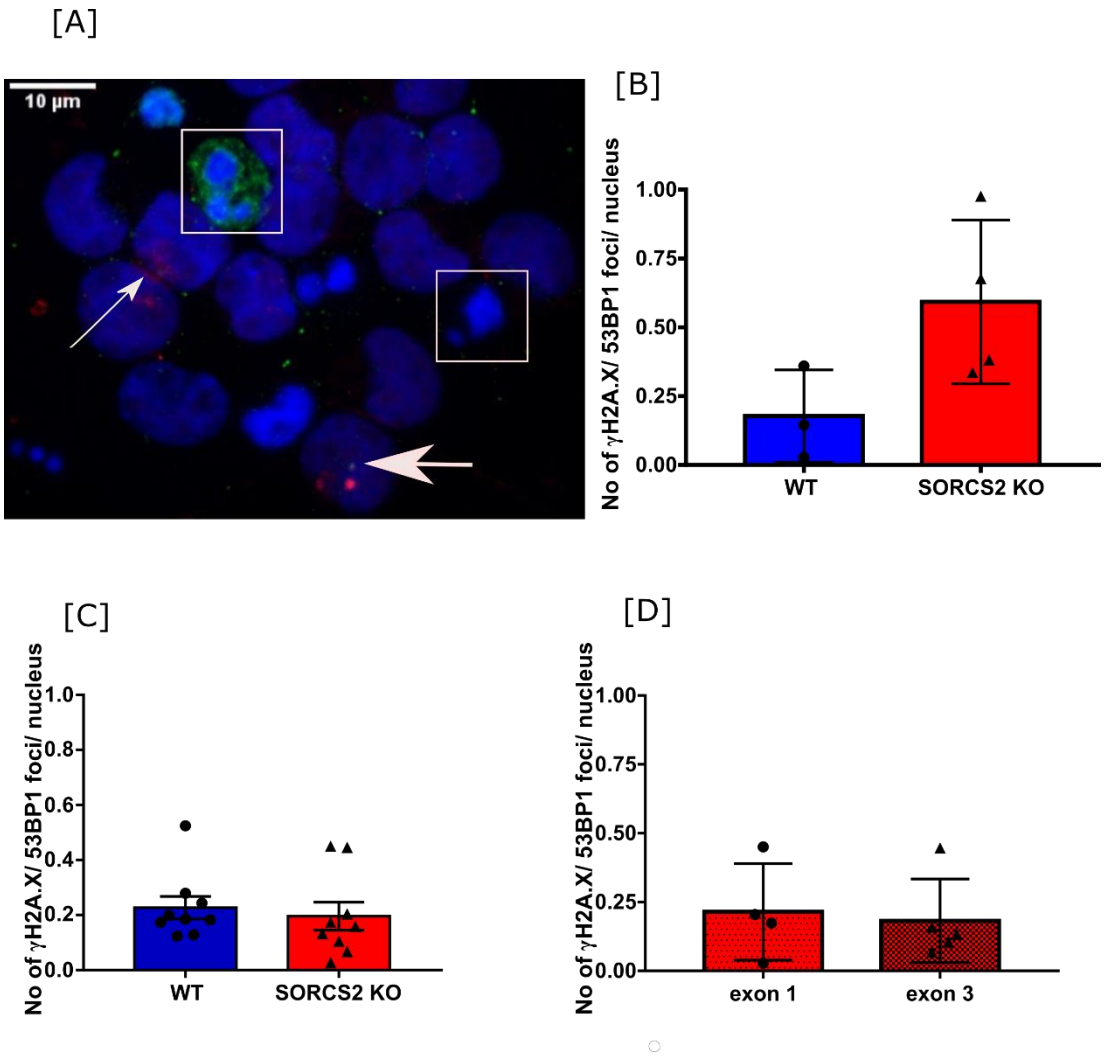

Figure 5.1. Levels of DNA DSB formation in untreated WT and SORCS2 KO LUHMES cell lines. [A] A representative z-stack confocal image (100x magnification). The thin arrow points at a nucleus (DAPI, blue) with diffuse 53BP1 staining (red), while the thick arrow points at a nucleus with a focus where γH2A.X (green) co-localised with 53BP1 (red) to give yellow staining. White rectangles show examples of nuclei excluded from the analysis due to either bright γH2A.X staining covering the whole nucleus or abnormal, apoptotic morphology. Scale bar corresponds to 10µm. [B] Initial experiment performed on a small number of WT (n=3) and SORCS2 KO (n=4) clones, comparing the number of γH2A.X/53BP1-positive foci per nucleus between the two genotypes. [C] The number of γH2A.X/53BP1-positive foci per nucleus in a follow-up experiment comparing WT LUHMES neurons (n=9) and SORCS2 KO clones (n=9) derived using two different gRNAs. [D] Comparison of the number of γH2A.X/53BP1-positive

foci per nucleus derived by targeting exon 1 or exon 3. Unpaired t-test, each point on the graph represents an individual clone; error bars represent mean ± SD.

5.2.2. Assessing the levels of DNA DSBs in the *SORCS2* KO LUHMES lines following treatment with etoposide

As described above, in post-mitotic cells, etoposide prevents the re-ligation of DNA DSBs occurring as a result of TopoIIβ activity[168]. Therefore, treatment with etoposide would lead to the accumulation of otherwise naturally forming, TopoIIβ-dependent DNA breaks, providing a 'snapshot' of all events that took place over the course of the treatment. In order to be able to both visualise and reliably count the number of DNA breaks following treatment with etoposide, I started by titrating the concentration of the compound used until single foci-like staining was observed. A range of etoposide concentrations, starting from 50μM to 0.5μM, were applied to WT LUHMES neurons (day 14) for 4 hours. The cells were then fixed and stained for γH2A.X. Foci-like staining pattern was observed after titrating the compound concentration down to 2μM. However, 0.5μM resulted in a clearer foci staining and thus all subsequent experiments involving treatment with etoposide were performed using this concentration (Figure 5.2.A).

Subsequently, I grew and differentiated into neurons WT (n=9) and *SORCS2* KO (n=9) clones of a similar passage number. At day 14 of differentiation, I treated the neurons with etoposide (0.5μM) for 4 hours. Approximately 100 nuclei (from four images taken from the same coverslip) were screened, and the number of foci positive for both γH2A.X and 53BP1 was counted for each cell line. Knocking out *SORCS2* led to a significant increase in the number of γH2A.X/53BP1-positive breaks per nucleus ($p= 0.0474$, Figure 5.2.B). The observed increase in DSB formation was unlikely to be the result of CRISPR off-target activity as there was no difference in the number of γH2A.X/53BP1-positive foci per nucleus between the *SORCS2* KO clones derived after targeting exon 1 and those generated after disrupting exon 3 ($p= 0.9322$, Figure 5.2.C).

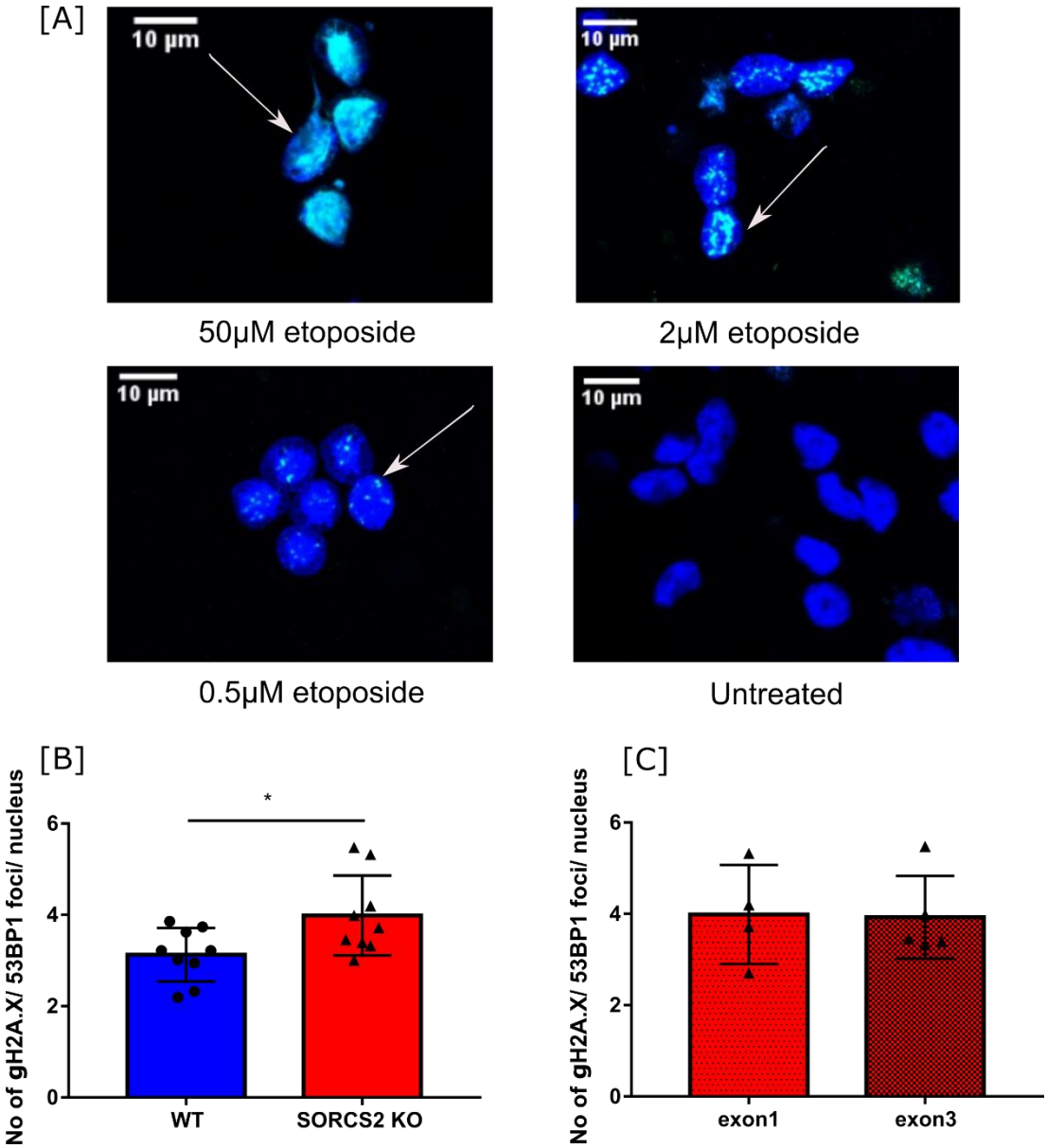

Figure 5.2. DNA DSB formation in WT and SORCS2 KO LUHMES neurons following treatment with etoposide. [A] Representative z-stack confocal images (60x magnification) of LUHMES WT neurons treated with various etoposide concentrations: 50µM (top left), 2 µM (top right) and 0.5 µM (bottom left), as well as untreated (bottom right). White arrows pointing at nuclei (DAPI, blue), where γH2A.X staining (green) was detected following treatment with etoposide. Scale bars correspond to 10µm. [B] Comparing the number of γH2A.X/53BP1-positive foci per nucleus in WT LUHMES neurons (n=9) and SORCS2 KO clones (n=9) following etoposide treatment. [C] Comparison of the number of γH2A.X/53BP1-positive foci per nucleus between etoposide-treated SORCS2 KO clones derived by targeting exon 1 (n=4) or exon 3 (n=5). Unpaired t-test, '' corresponds to p<0.05; each dot on the graphs corresponds to an individual cell line; error bars represent mean ± SD.*

5.2.3. Assessing neuronal viability in the *SORCS2* KO clones

Since DSB formation has been shown to affect neuronal function and survival[169], I wanted to see whether knocking out *SORCS2* has an impact on the overall neuronal viability. The same WT (n=9) and *SORCS2* KO lines (n=9) used in the experiments described above were brought up and differentiated into neurons. Viability was measured at day 6 and day 14 of differentiation using the Alamar Blue assay, which provides a measure of the metabolic activity of the cells. At day 6, there was no significant difference in the viability of WT (n=9) neurons compared to that of *SORCS2* KO clones (n=9) (p=0.7144; Figure 5.3.A). However, at day 14, *SORCS2* KO clones (n=7) showed a significant reduction in their viability compared to WT controls (n=9) (p=0.0047, Figure 5.3.B). All the cells in two *SORCS2* KO lines lifted off the plate prior to day 14, preventing assessment of viability.

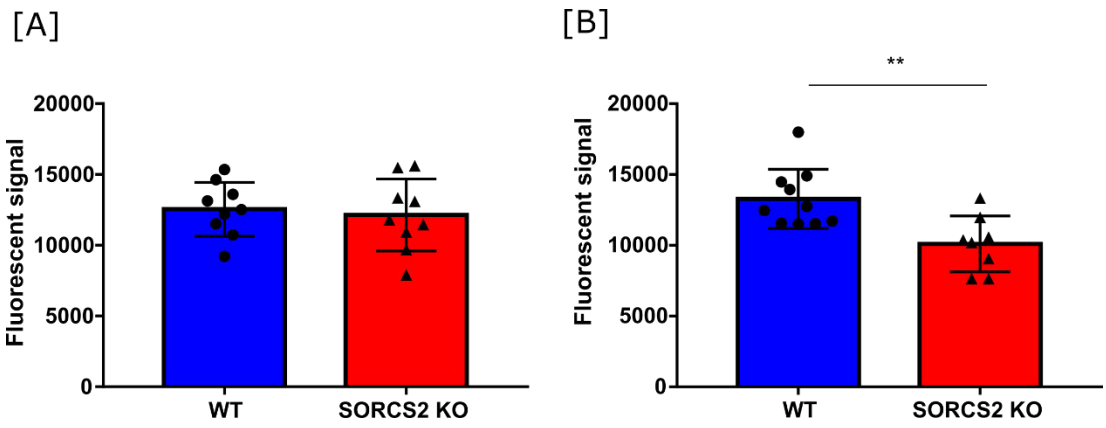

*Figure 5.3. Assessing neuronal viability, measured as fluorescent signal, in the SORCS2 KO clones using the Alamar Blue assay. [A] Comparison of neuronal viability between WT (n=9) and SORCS2 KO (n=9) cell lines at day 6 of differentiation. [B] Comparing neuronal viability of WT (n=9) and SORCS2 KO (n=7) at day 14 of differentiation. Unpaired t-test; each dot on the graph represents an individual cell line; '**' stands for p< 0.01.*

5.3. Investigating potential mechanisms underlying the increase in DNA DSBs upon *SORCS2* knockout

5.3.1. Assessing the effect of knocking out *SORCS2* on DNA DSB formation following NMDA stimulation *in vitro*

NMDA is a potent NMDAR agonist, and non-cytotoxic NMDAR stimulation with this compound has been previously associated with increased γH2A.X foci formation in primary neuronal cultures[107,115]. Moreover, Madabhushi *et al.* (2015)[118] showed that these foci were located primarily within early-response genes, similar to those detected following etoposide treatment. In addition, culturing primary neurons under conditions favouring the stimulation of either NR2B- or NR2A-containing NMDAR, implicated the former in the formation and the latter in the repair of DNA DSBs. Given these findings, as well as the role of SorCS2 in NMDA receptor trafficking[51], I hypothesised that the *SORCS2* KO LUHMES lines may exhibit a different response in terms of DNA DSB formation upon stimulation with NMDA compared to control lines.

To address this hypothesis, first I determined the optimal NMDA concentration and treatment length that led to the highest increase in DSBs in WT neurons (day 14) compared to untreated controls. I used two previously established protocols for treating primary neurons with NMDA as a guidance. In one of the protocols, Crowe *et al*. (2006)[115] applied different concentrations of NMDA (15µM and 50µM) and fixed the neurons after various exposure times, ranging from 10 minutes to 4 hours. They observed the highest increase in the number of DSBs in the neurons treated with 15µM of NMDA for 10min[115]. In another study, Suberbielle *et al.* (2013)[107] reported maximum effect after five-minute treatment with 50µM NMDA, followed by 1 hour recovery. In both studies the number of DSBs returned to baseline levels after 4 hours of drug application. In order to find the NMDA concentration, application procedure and time that worked best in LUHMES, I tested the same conditions as the ones described by Crowe *et al.* (2006)[115] and Suberbielle *et al.* (2013)[107]. In day 14 LUHMES neurons, the strongest effect in terms of γH2A.X foci formation was observed after applying 50µM NMDA for 5 minutes, followed by 1 hour of recovery in differentiation media (Figure 5.4.A).

Subsequent experiments performed on multiple day 14 WT LUHMES lines (n=3), stained for both γH2A.X and 53BP1, showed a significant increase in the number of DSBs per nucleus using these treatment conditions (p=0.0003, Figure 5.4.B). In parallel, *SORCS2* KO lines (n=3), in which exon 3 was targeted, were also brought up and differentiated for 14 days. Statistical analysis showed no difference in the degree of NMDA-induced DNA DSB formation, examined as the average number of γH2A.X/53BP1 foci per nucleus, in the *SORCS2* KO clones compared to treated WT neurons of a similar passage number (p=0.2668, Figure 5.4.C). Due to issues caused by a change in the cell culture plastics and time limitations, I was unable to repeat this analysis on a larger number of cell lines.

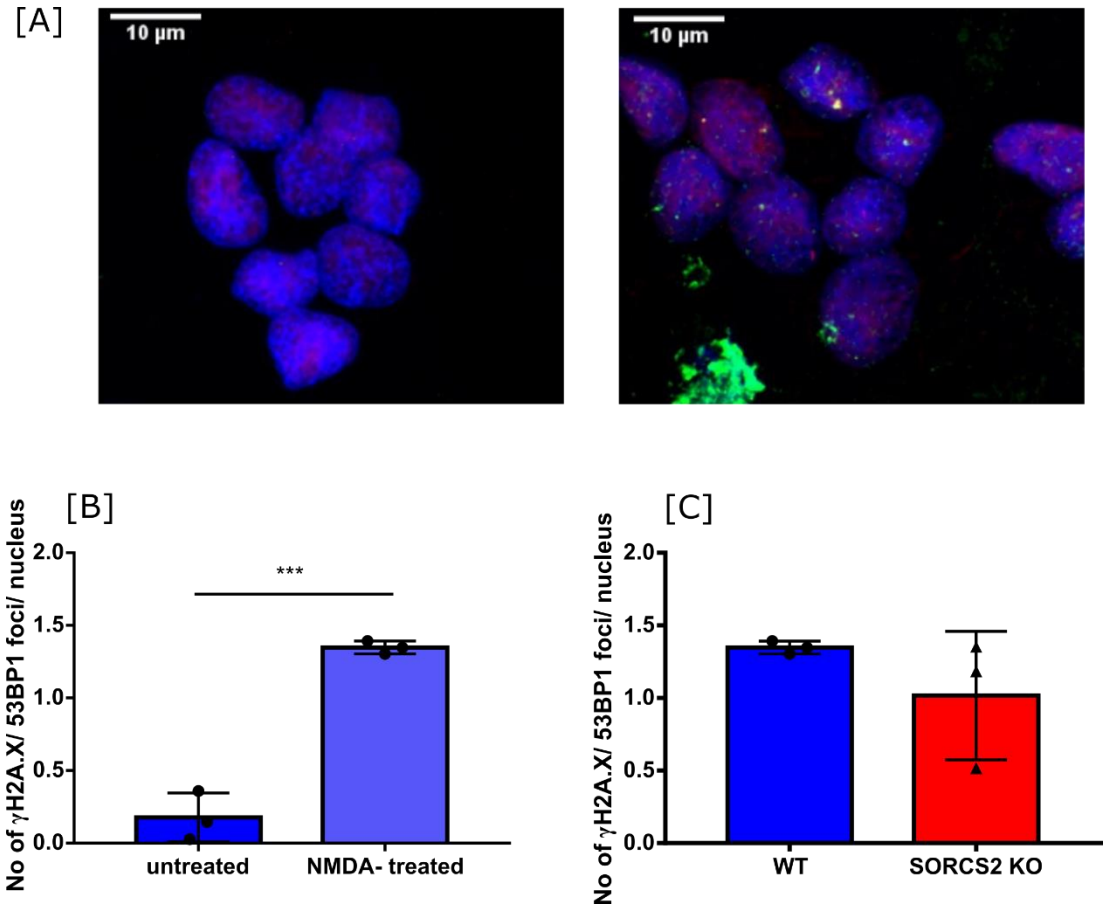

Figure 5.4. DNA DSB formation in WT and SORCS2 KO LUHMES neurons following treatment with NMDA. [A] Representative z-stack confocal images (100x magnification) comparing untreated (left) LUHMES WT neurons with neurons treated with NMDA (right). Scale bars correspond to 10µm. [B] Comparison of the number of γH2A.X/53BP1-positive foci per nucleus in untreated and NMDA-treated WT LUHMES neurons (day 14). [C] Comparison of the number

*of γH2A.X/53BP1-positive foci per nucleus in WT LUHMES neurons (n=3) and SORCS2 KO clones (n=3) following NMDA treatment. Unpaired t-test, '***' stands for p<0.001; each dot on the graphs represents and individual cell line; error bars represent mean ± SD.*

5.3.2. Assessing extracellular Aβ levels in the *SORCS2* KO LUHMES lines as a potential cause of DNA DSBs

As discussed in the introduction, elevated levels of Aβ have been linked to epileptiform neuronal activity and increased DNA DSB formation both *in vivo* and *in vitro*[107]. Additionally, SORCS2, as all other members of the sortilin receptor family, has been implicated in APP processing and thus Aβ production[46]. Therefore, I hypothesised that increased Aβ levels might contribute to the exacerbated DNA DSB formation observed in the *SORCS2* KO LUHMES lines following treatment with etoposide, and thus potentially in the *Sorcs2*$^{-/-}$ mice.

To address this hypothesis, I decided to compare the levels of secreted $A\beta_{42}$ between WT and *SORCS2* KO clones. Initially, I differentiated *SORCS2* KO clones generated by targeting exon 3 (n=3), together with WT lines (n=3). One of the WT lines died during the differentiation process. At day 14 of differentiation, I collected media from the *SORCS2* KO (n=3) and the remaining WT (n=2) lines and performed $A\beta_{42}$ sandwich ELISA. However, the obtained results were below the assay detection limit of 15.6pg/ml (Table 5.1.).

Serine proteases have been shown to rapidly degrade Aβ peptides in cell culture media[170], and might have thus explained the low levels of extracellular $A\beta_{42}$ detected in the collected samples. To account for this, as well as for the small sample size of the initial experiment, I cultured six WT lines and nine *SORCS2* KO clones- four generated by targeting exon 1, and five obtained after disrupting exon 3. As before, I collected media from these lines following 14 days of differentiation. This time I added a protease inhibitor cocktail, containing the serine protease inhibitor AEBSF, to each of the medium samples at a final concentration of 1mM. Nonetheless, the amount of extracellular $A\beta_{42}$ detected in the collected samples using the same ELISA kit was still below the assay detection range for all samples (Table 5.1.). Therefore, the results were considered unreliable and discarded.

Table 5.1. A table summarising the concentration of extracellular $A\beta_{42}$ (pg/ml) obtained for each of the tested WT and SORCS2 KO cell lines from the two ELISA experiments. All samples measured were below the assay lowest detection limit of 15.6pg/ml.

Cell line	**Extracellular $A\beta_{42}$ concentration (pg/ml)**
First ELISA experiment	
WT1	3.63
WT2	2.45
***SORCS2* exon 3 clone 36**	3.73
***SORCS2* exon 3 clone 39**	3.04
***SORCS2* exon 3 clone 40**	2.84
Second ELISA experiment	
WT1	3.65
WT2	4.14
WT3	4.81
WT 4	2.86
WT 5	2.80
WT 6	2.50
***SORCS2* exon 1 clone 5**	2.98
***SORCS2* exon 1 clone 44**	3.04
***SORCS2* exon 1 clone 51**	3.1
***SORCS2* exon 1 clone 59**	3.35
***SORCS2* exon 3 clone 26**	3.77
***SORCS2* exon 3 clone 36**	2.92
***SORCS2* exon 3 clone 39**	3.35
***SORCS2* exon 3 clone 40**	2.92
***SORCS2* exon 3 clone 41**	3.35

5.3.3. Assessing extracellular dopamine levels in the *SORCS2* KO LUHMES lines

In mice, SorCS2 plays an important role in the development of the dopaminergic system[6]. Knocking out *Sorc2* leads to increased dopaminergic innervation from the VTA to the frontal cortex, accompanied by reduced dopamine levels as a result of altered dopamine metabolism in this brain area[6]. In addition, both dopamine and its metabolite L-DOPA have been shown to induce DNA damage, via their interaction with copper ions[164], and endogenous dopamine has been shown to play a role in controlling dopamine neuron firing patterns[171]. Given these findings, the role of TopoIIβ in gene expression following neuronal activity[118] and the fact that the LUHMES cell line constitutes a homogenous population of dopaminergic neurons, I decided to test the hypothesis that altered dopamine secretion might contribute to the increase in DNA DSB formation observed in the *SORCS2* KO cell lines following treatment with etoposide.

Abbie Payne (an MSc student in our lab) grew and differentiated WT (n=6) and *SORCS2* KO (n=9) cell lines of a similar passage number under the supervision of myself and Susan Anderson. As described in the above sections, both *SORCS2* KO clones generated after targeting exon 1 and exon 3 were included to account for CRISPR off-target activity. At day 14, I collected media from each line and subsequently measured the amount of extracellular dopamine using a competitive dopamine ELISA kit. The amount of dopamine detected was normalised to the total amount of protein collected from the corresponding cell line and measured by a BCA assay. There was no significant difference in the amount of dopamine released by the *SORCS2* KO clones when compared to the WTs (p= 0.0857, Figure 5.5).

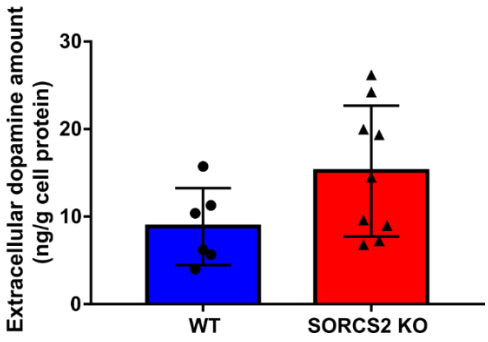

Figure 5.5. Assessing extracellular dopamine levels in the generated SORCS2 KO LUHMES clones. Extracellular dopamine levels (ng) normalised to the total protein among (g) in

SORCS2 KO (n=9) compared to WT (n=6) day 14 LUHMES neurons. Unpaired t-test, each dot represents an individual cell line; error bars represents mean ± SD.

5.4. Assessing the levels of DNA DSB formation in *Sorcs2*$^{-/-}$ mouse primary neurons

5.4.1. Assessing the baseline levels of DNA DSB formation in *Sorcs2*$^{-/-}$ primary neurons

Since I detected elevated levels of DNA DBS formation in the dentate gyrus of *Sorcs2*$^{-/-}$ mice, I decided to investigate whether I would observe the same increase in primary hippocampal neurons derived from these mice. All experiments described below were performed in the lab of our collaborator, Assoc. Prof Simon Glerup, Aarhus University, Denmark.

Hippocampal primary neurons were obtained from *Sorcs2*$^{-/-}$ and WT pups (P0) and fixed after being maintained for 12 days *in vitro* (DIV12). Three independent cultures per genotype were stained for γH2A.X and 53BP1, and images were obtained using confocal microscope. As before, approximately 100 nuclei were counted, and nuclei showing apoptotic morphology or being fully covered with bright γH2A.X signal were excluded from the analysis. The number of γH2A.X/53BP1-positive foci per nucleus was calculated. There was no significant difference in the number of foci per nucleus when *Sorcs2*$^{-/-}$ neurons were compared to WT controls (p= 0.0580; Figure 5.6.B).

5.4.2. Assessing the levels of DNA DSBs in *Sorcs2*$^{-/-}$ primary hippocampal neurons following treatment with etoposide

Parallel to the experiment in the untreated neurons, three independent cultures of *Sorcs2*$^{-/-}$ and WT primary neurons (DIV 12) were treated with etoposide (0.5μM) for 4 hours. Neurons were then fixed and stained for γH2A.X and 53BP1. Approximately 100 nuclei were counted, and nuclei showing apoptotic morphology or being fully covered with bright γH2A.X signal were excluded from the analysis. As for the untreated cells, there was no significant difference in the number of γH2A.X/53BP1-

positive foci per nucleus between the *Sorcs2*$^{-/-}$ and WT primary neurons (p=0.4238; Figure 5.6.C).

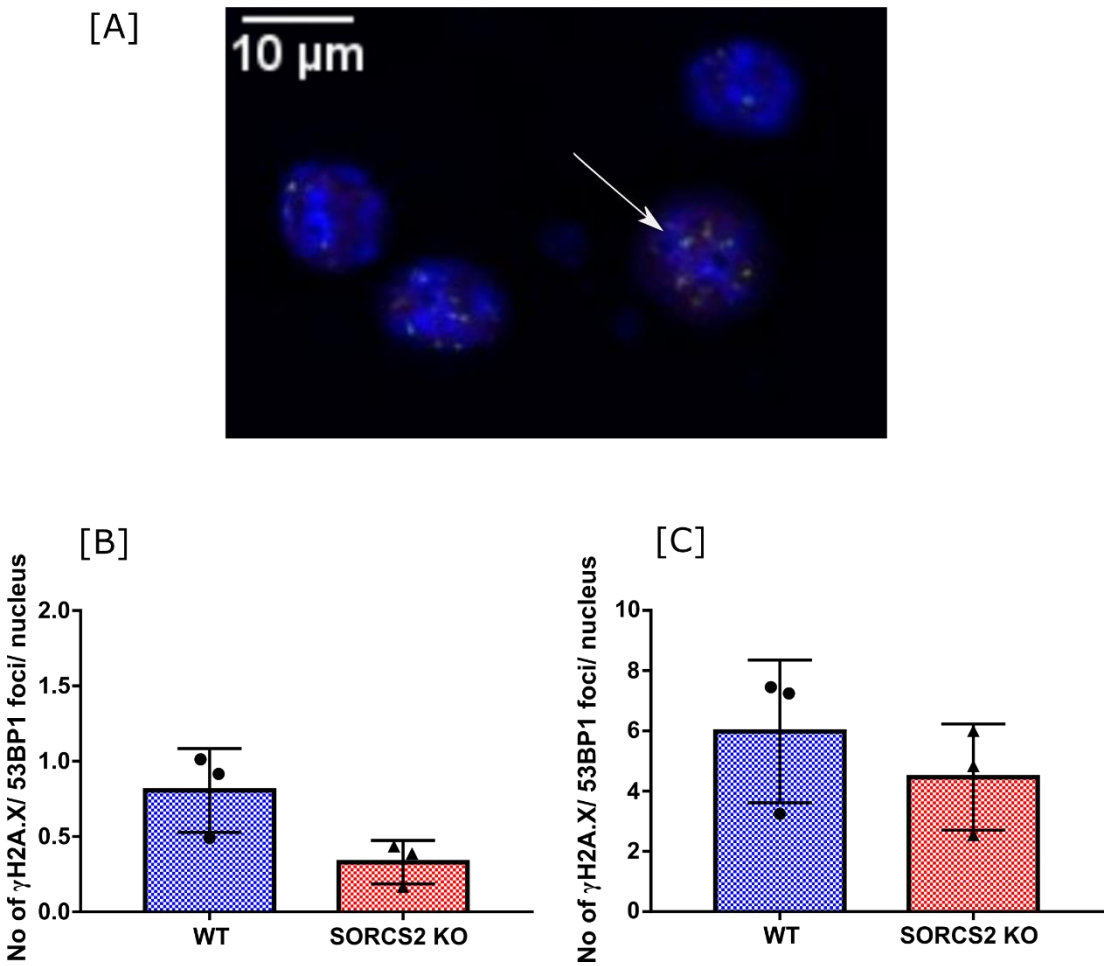

Figure 5.6. Levels of DNA DSB formation in the Sorcs2$^{-/-}$ *and WT primary neurons (DIV12) with and without treatment with etoposide. [A] A representative z-stack confocal image (60x magnification) of etoposide-treated cells. White arrow points at a nucleus (DAPI, blue) with foci positive for γH2A.X (green) and 53BP1 (red). Scale bar corresponds to 10µm. [B] Comparison of the number of γH2A.X/53BP1-positive foci per nucleus between untreated Sorcs2*$^{-/-}$ *and WT primary neurons (DIV12) from three independent cultures. [C] Comparison of the number of γH2A.X/53BP1-positive foci per nucleus in etoposide-treated Sorcs2*$^{-/-}$ *and WT primary neurons (DIV12) from three independent cultures. Unpaired t-test, each dot on the graph represents an independent culture; error bars represent mean ± SD.*

5.5. Discussion

Suberbielle *et al.* (2013)[107] reported a transient increase in DSB formation (detected as formation of γH2A.X foci) in the neurons of WT mice, subjected to a learning and memory task, i.e. exploration of a novel environment. They linked the observed phenotype to physiological increase in neuronal activity. More specifically, they implicated NR2B-containing NMDARs in the formation, and NR2A-containing NMDARs in the repair of these DNA breaks. In addition, hAPP mice showed increased baseline levels of neuronal DNA DSBs, as well as compromised repair following exposure to the same novel environment. The authors attributed these differences to the aberrant network activity previously reported for hAPP mice. Further *in vitro* experiments showed that the observed Aβ-induced increase in DNA DSB formation is the result of altered balance between NR2A- and NR2B-containing NMDAR activity[107].

Similarly to Suberbielle *et al.*, Madabhushi *et al.* (2015)[118] detected increased γH2A.X foci formation as a result of neuronal activity *in vivo* (in a fear conditioning task), as well as in *ex vivo* hippocampal slices stimulated with NMDA or theta bursts and in primary neurons treated with KCl, bicucculin or NMDA. Stimulation of primary neurons with these compounds was also associated with increased expression of early-response genes (e.g. *Npas4* and *Fos*), and ChiP sequencing with γH2A.X confirmed that the detected DSBs lay in regions of the genome corresponding to these genes. Treatment with etoposide (TopoIIβ-specific poison in post-mitotic cells) upregulated the expression of the same set of early-response genes. Further experiments showed that the formation of DNA DSBs induced by neuronal activity occurred within promoter regions of early-response genes and was the result of TopoIIβ activity.

As discussed in chapter 3, I detected increased levels of DSB formation in *Sorcs2*$^{-/-}$ mice compared to WT controls. Subsequently, as described in chapter 4, I knocked out *SORCS2* in LUHMES cells in order to test whether the results from the mice could be replicated in human neurons and if so, to explore potential mechanisms that might explain the phenotype. As described in this chapter, first I examined the levels of DNA DSB formation in untreated *SORCS2* KO clones by counting the number of foci where γH2A.X co-localised with the DDR protein 53BP1 per nucleus. Initial experiments performed on a small number of WT (n=3) and *SORCS2* KO (n=4) lines showed no significant difference in the number of DSBs per nucleus and a high degree of inter-clonal variation for both genotypes. When the experiment was repeated on a larger

sample size, including additional WT and *SORCS2* KO clones in addition to the ones tested in the first experiment, the lack of significant difference between the two genotypes was confirmed. Overall, the levels of variation in this experiment were lower compared to the first experiment, especially in the control group. There was no difference in the number of DNA breaks per nucleus between the *SORCS2* KO clones generated by targeting exon 1 compared to the KO clones obtained by disrupting exon 3, which justified considering them as one group.

Meanwhile, as previously reported, treating WT LUHMES neurons with etoposide led to a significant increase in the number of γH2A.X/53BP1-positive foci per nucleus. Moreover, *SORCS2* KO clones (n=9) showed a significant increase in the number of γH2A.X/53BP1-positive foci per nucleus compared to WT lines (n=9) grown simultaneously. It is unlikely that the observed increase was the result of CRISPR off-target activity as *SORCS2* KO clones generated with two different gRNAs targeting exon 1 or exon 3 showed comparable levels of DNA damage.

The lack of significant difference in the number of DSBs per nucleus between untreated WT and *SORCS2* KO clones could be explained with the low and highly variable levels of DNA damage observed in the untreated cells, and especially in the *SORCS2* KO clones. Meanwhile, treatment with etoposide leads to the accumulation of TopoIIβ-dependent DSBs, providing a more reliable measure of all events that took place over the course of the treatment. Nonetheless, the fact that I was able to detect significant difference in the number of DSBs after treatment with etoposide, but not in the untreated cells, may suggest that knocking out *SORCS2 in vitro* exacerbates TopoIIβ-dependent DNA DSB formation but has no effect on their repair. However, the experiment will need to be repeated for these results to be confirmed. Unfortunately, due to changes in the cell culture plastics and time constraints, I was unable to do so.

Alongside the increased DSB formation observed in the etoposide-treated *SORCS2* KO neurons, I also detected reduced viability in these clones at day 14 but not day 6 of differentiation. These results compliment a recently described role of SORCS2 in neuroprotection[22]. Moreover, the fact that the difference in the viability was observed at day 14 but not day 6 coincides with *SORCS2* being expressed at low levels in LUHMES cells until after day 10 of differentiation, as described in chapter 4. Accumulation of DNA DSBs has been shown to affect neuronal function and survival. However, given that I did not observe any difference in the number of DSBs per

nucleus in the untreated *SORCS2* KO versus WT clones, it is unlikely that the reduced viability detected is the result of accumulated DNA damage.

As discussed above, neuronal activity triggers TopoIIβ-dependent DSB formation. Therefore, I hypothesised that altered neuronal activity could explain the elevated levels of DNA DSBs in the *Sorcs2*$^{-/-}$ mice and in the etoposide-treated *SORCS2* KO LUHMES lines. As discussed above, increased Aβ levels and altered balance between NR2A- and NR2B-containing NMDAR activity have been associated with aberrant neuronal activity and exacerbated DNA DSB formation[107]. As in previously reported experiments in mouse primary neurons[107,115], stimulating WT LUHMES neurons (day 14) with NMDA increased significantly the number of γH2A.X/53BP1-positive foci per nucleus. SorCS2 is involved in the trafficking of NMDAR subunit NR2A[51], which, as mentioned above, has been implicated in the repair of DNA DSBs[107]. I did not observe any difference in the extent of NMDA-induced DNA DSB formation in the *SORCS2* KO clones compared to WT controls. However, this experiment was performed on a small number of clones and thus may not reflect a real lack of difference. Therefore, further experiments including a larger number of cell lines are required to assess the effect of knocking out *SORCS2* on NMDA-induced DSB formation. Unfortunately, due to changes in cell culture plastics and time limitations, I was not able to repeat this experiment on a larger number of cell lines. However, if these preliminary results were confirmed in subsequent experiments, this would suggest that the elevated levels of DNA breaks observed in the etoposide-treated *SORCS2* KO LUHMES neurons are not the result of dysregulated NMDAR signalling.

Previous work in HEK293 cells demonstrated that knocking down *SORCS2* promotes APP amyloidogenic processing[46]. Unfortunately, I was unable to detect extracellular $Aβ_{42}$ in the media collected from both WT and *SORCS2* KO LUHMES neurons. This did not allow to rule out the possibility that elevated Aβ levels might contribute to the increased levels of DNA DSBs detected in the etoposide-treated knockout lines. Both $Aβ_{40}$ and sAPPβ have been previously detected in mature LUHMES neurons[133,134]. Thus, it would be interesting to see whether knocking out *SORCS2* in LUHMES exerts an effect on the production of these APP metabolites. Alternatively, a more sensitive $Aβ_{42}$ assay could be used to try and detect differences in this particular Aβ species.

Both untreated and treated *Sorcs2*$^{-/-}$ primary neurons showed no significant difference in the number of DNA DSBs per nucleus compared to WT controls. Major caveats of

this experiment were the small sample size and the high degree of variation between the different cultures, questioning whether the obtained results reflect a true lack of difference or the collected data lacks power to detect a difference between the two groups. Ideally, to overcome these limitations, the analysis would need to be performed on a larger number of independent cultures. Unfortunately, due to time limitation and problem with mice breeding, I was unable to obtain more than three cultures for each genotype. However, it should be noted that while not statistically significantly different, the absolute levels of DSBs observed in the *Sorcs2*$^{-/-}$ neurons is lower compared to the WTs. This is in contrast with the results obtained in the LUHMES neurons. If this difference was to be maintained on repetition with a larger dataset, possible explanation would include the origin of the neurons (i.e. the primary neurons were isolated from mice, while LUHMES are human neurons), as well as their type (hippocampal versus dopaminergic neurons) and electrophysiological properties. LUHMES neurons have been shown to fire spontaneous action potentials after day 10 of differentiation[133]. However, their electrophysiological properties following 14 days of differentiation cannot be directly compared to these of primary hippocampal neurons derived from new-born (P0) mice and maintained in culture for 12 days.

As mentioned above a possible interpretation of the results obtained might be that the observed elevated levels of etoposide-induced DNA breaks in the LUHMES KO lines are the result of altered mechanisms, specific to dopaminergic neurons. Endogenously released dopamine is involved in the regulation of dopamine neuron firing patterns *in vivo* by acting on the dopamine autoreceptor D(2)R[171]. Moreover, increased dopamine levels have been previously linked to DSB formation[164]. Despite its role in the development of the dopaminergic system, knocking out *SORCS2* in LUHMES had no effect on the amount of dopamine secreted by these cells. However, the *SORCS2* KO clones displayed much higher levels of variation in the amount of secreted dopamine than WT neurons. Nonetheless, based on these results, it is unlikely that altered extracellular dopamine levels could explain either the increased levels of etoposide-induced DNA DSBs in these cells or the discrepancy with the results obtained in the primary neurons.

Chapter 6 Assessing the effect of introducing the rare EOAD-associated mutation G508S on *SORL1* expression and Aβ production

6.1. Introduction

SORLA plays an important role in APP processing and thus Aβ production. APP amyloidogenic processing involves the endocytosis of APP molecules and their transport to late endosomes, where proteolytic breakdown of APP by β- and γ-secretase results in the formation of Aβ[172]. SORLA has been shown to act as an APP sorting receptor, shuttling APP molecules from endosomes to the trans-Golgi network and thus preventing their entry into both amyloidogenic and non-amyloidogenic processing pathways[37]. In addition to its interaction with APP, SORLA also impacts on the amyloidogenic burden by direct binding of its Vps10p domain to Aβ[41]. *In vitro*, the association between SORLA and Aβ leads to increased lysosomal targeting and degradation of the Aβ peptides[41]. In keeping with these findings, loss of SorLA in mice accelerates Aβ production and senile plaque deposition[37, 42, 43].

Reduced SORLA levels have been described in post-mortem brains from individuals with AD[88, 39]. Moreover, GWASs of AD, including the most recent meta-analysis, have identified significant association with *SORL1* variants[65,68,69]. The associated variants are largely non-coding and are expected to exert their effects via altering expression levels. However, rare *SORL1* coding variants, including missense, splice site and nonsense mutations, have also been linked to both familial and sporadic AD and both early and late onset forms of the disease[41,73,76,77,80,173–176] . Functional studies have been performed to assess the impact of some of the identified missense mutations[41,76]. For example, the G511R missense mutation, associated with familial AD, was shown to disrupt SORLA binding to Aβ, resulting in decreased Aβ turnover[41]. Overexpression of five other *SORL1* mutations: T588I and T2134M (found in EOAD cases)[76] and E270K, A528T and T947M (found in LOAD cases)[75] reduced SORLA interaction with APP and led to increased APP proteolytic processing and Aβ secretion. Additionally, two silent mutations located at the 3' region of the gene were shown to decrease SORLA protein levels in post-mortem AD brains, as well as to reduce translation efficiency *in vitro*[177]. Nevertheless, the functional consequences of the majority of the *SORL1* coding variants identified so far remains unknown.

In summary, there is strong evidence from multiple data types linking SORLA to AD. While the cellular mechanisms are not yet fully elucidated, it appears that decrease or loss of *Sorl1* expression in animal models, as well as some missense mutations in SORLA, lead to increased levels of Aβ. Therefore, it can be hypothesised that knocking out *SORL1* in human neurons or introducing *SORL1* variants associated with EOAD would lead to an increase in Aβ production, and thus in DNA DSBs, as has been previously observed in hAPP mice and primary neurons treated with Aβ oligos[107].

6.1.1. Aims

The aims of the work described in this chapter were to knock out *SORL1* and to introduce the rare EOAD-associated missense mutation G508S into LUHMES cells using CRISPR/Cas9 genome editing. A further aim was to examine the impact of the mutations on *SORL1* expression and Aβ production.

6.2. Knocking out *SORL1* in LUHMES cells using CRISPR/Cas9 genome editing

6.2.1. First attempt to knock out *SORL1* in LUHMES cells using CRISPR/Cas9 genome editing

SORL1 is located on chromosome 11 and encodes for a large number of transcripts, the majority of which start after exon 25 of the full-length, protein-coding transcript. Previous work in the lab performed by Susan Anderson has shown that in addition to the canonical full-length isoform, LUHMES express several shorter isoforms whose 5' ends lie within the second half of the 48 exons. Therefore, I decided to knock out the gene by targeting the first exon common for all expressed *SORL1* transcripts- exon 31. I transfected proliferating cells of a low passage number (P5) with the pSpCas9(BB)-2A-GFP plasmid containing a gRNA, targeting this exon. The transfection efficiency was low- approximately 1%, and the top 3% of GFP-positive cells were FAC sorted into 96-well plates. Ten days post-transfection, I identified seven colonies derived from a single cell. These were then transferred into 48-well plates and further expanded to 6-well plates. One clone (1.4) did not survive the

transfer process. Once the clones reached 85% confluency in the 6-well plates, frozen stocks were generated, and proportion of the cells were kept for DNA isolation and genotyping.

6.2.1.1. Genotyping of the identified clonal lines

DNA was isolated and PCR performed using primers flanking the targeted *SORL1* exon 31. Three of the isolated clonal lines showed a single band on the gel, three others had two bands and one was discarded as it had three (Figure 6.1.A). Where two bands were present on the gel (clones 1.2, 1.3 and 1.6), TOPO cloning was performed, and the resulting products were subjected to sequence analysis, the results of which are summarised in Table 6.1.

Table 6.1. A table listing all clones isolated following targeting of SORL1 exon 31, the PCR and sequencing results, and the consequences of the mutations introduced.

Clone	PCR	Sequencing	Impact
1.1	Three bands	Not sequenced, discarded	
1.2	Double band	Allele 1: 1bp insertion Allele 2: 11bp insertion	Frameshift mutation
1.3	Double band	Allele 1: 2bp insertion Allele 2: 11bp insertion	Frameshift mutation
1.4	N/A- clone died		
1.5	Single band	11bp deletion	Frameshift mutation
1.6	Double band	Allele 1: multiple deletions Allele 2: WT	Frameshift mutation
1.7	Single band	6bp deletion	Inframe deletion

6.2.1.2. Assessing *SORL1* mRNA levels in the obtained *SORL1* mutant clones

In order to assess the expression of *SORL1* in the CRISPR-edited clones identified, I performed qRT-PCR on cDNA derived from proliferating cells, using two TaqMan probes, located near the 5' (*SORL1* exon 2-3) and the 3' end (*SORL1* exon 46-47) of the full-length *SORL1* transcript. The former aimed to detect expression levels of the full-length *SORL1* protein-coding isoform, while the latter would also amplify additional protein-coding transcripts, found to be expressed in LUHMES. *SORL1*

mRNA levels were normalised to the expression levels of two reference genes (*CYC1* and *TBP*), identified from a panel of eight as the most stable in this sample set by the GeNorm software (data not shown). All clones, apart from 1.5, showed lower *SORL1* mRNA levels compared to a WT LUHMES line grown simultaneously (Figure 6.1.B).

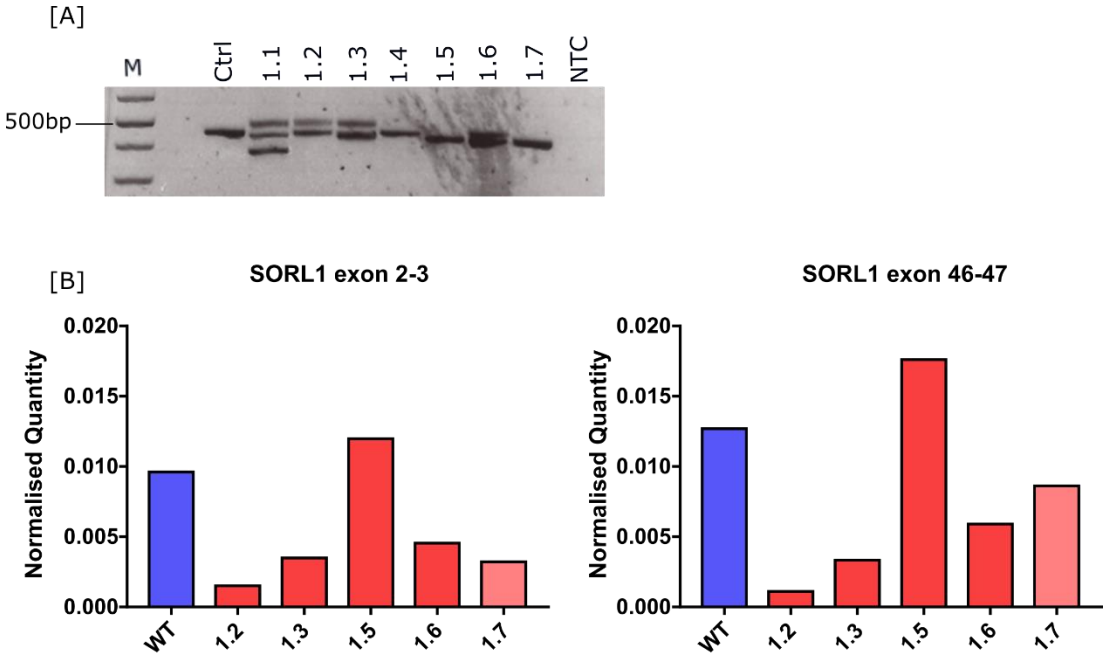

Figure 6.1. Results of targeting SORL1 exon 31 using CRISPR/Cas9 genome editing. [A] Genotyping of the isolated CRISPR-edited clones. PCR with primers, flanking SORL1 exon 31, on genomic DNA isolated from a WT LUHMES line ('Ctrl') and the seven clones identified after FACS (1.1- 1.7), run on 2% LMP gel loaded with 1 Plus kb ladder and stained with SYBER Safe. The first track was loaded with the ladder ('M'), the second was left empty, the third track was loaded with PCR product amplified from WT LUHMES DNA ('Ctrl') and the rest with samples from the seven isolated clones. 'NTC' stands for non-template control. Expected PCR product size- 462bp. [B] qRT-PCR assessing the SORL1 mRNA levels in the isolated clones, excluding clones 1.1 (which was discarded as it showed three bands on the gel) and 1.4 (which died during culturing). The x-axis shows the clones, while the y-axis- SORL1 mRNA levels normalised to the geometric mean of CYC1 and TBP. The blue bar corresponds to the WT line, while the red and the pink bars to clones, carrying frameshift or inframe mutations, respectively. Graphs present the results from a single experiment, and bars correspond to the average of three technical replicates.

6.2.2. Second attempt to knock out *SORL1* in LUHMES cells

The first attempt of knocking out *SORL1* resulted in four clones with *SORL1* mRNA levels lower than the WT. Amongst these, clones 1.2 and 1.3 showed the lowest expression (approximately 10% and 30% of the WT expression levels, respectively), suggesting they might constitute *SORL1* knockouts. However, due to both the low number of surviving clones, as well as of potential KO clones, I decided to repeat the experiment using the same gRNA. This time the transfection efficiency was 8.22%. Twenty-six single-cell colonies were identified ten days post-FACS and transferred to 48-well plates for further expansion, with proportion of cells from each clone kept for genotyping. One clone (2.24) did not survive the process of transfer, and another three (2.16, 2.23 and 2.25) died during expansion. Subsequently, I extracted DNA by boiling the cells in Tris buffer and performed PCR amplification, using the same primers as above. Amongst the identified clones, 14 showed single bands on the PCR gel, eight- double, and one was discarded as it had three bands (Figure 6.2.A). The DNA extracted from two clones failed to give a PCR product (Figure 6.2.A). Sequencing analysis of the clones with single bands identified two clones as CRISPR WTs, and the rest as carrying indels of different sizes (Table 6.2). Since mRNA levels do not necessarily correspond to protein, I screened the clones using western blot analysis. I selected one CRISPR WT (clone 2.9) and two clones predicted to carry frameshift mutations and thus potentially constituting *SORL1* KOs (clones 2.11 and 2.12) and differentiated them into neurons, alongside a LUHMES WT cell line of a similar passage number. After 14 days of differentiation, I collected protein samples from these cell lines and subjected them to western blot analysis. Neither of the two frameshift mutants showed changes in SORLA protein levels compared to the LUHMES WT line grown simultaneously (Figure 6.2.B).

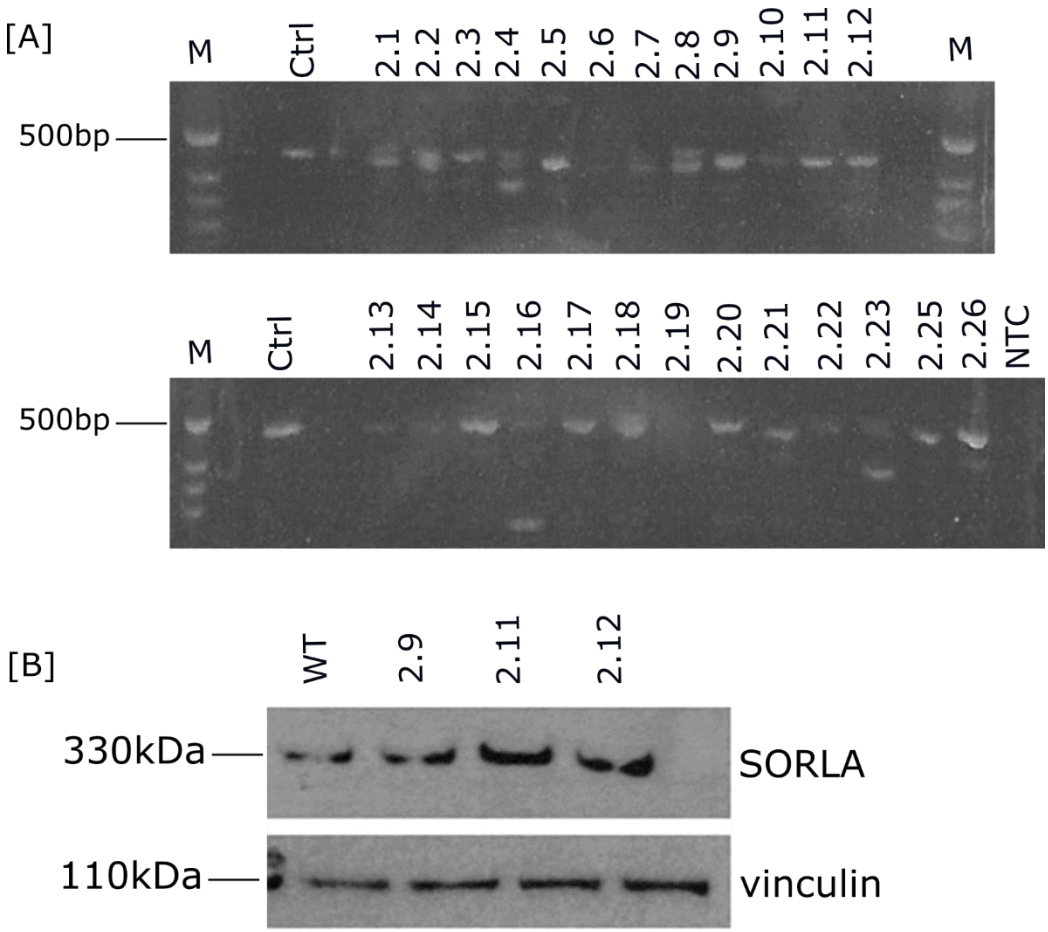

Figure 6.2. Results of the second attempt to knock out SORL1 by targeting exon 31. [A] Genotyping of the isolated single cell colonies. PCR with primers, flanking SORL1 exon 31, run on 1.7% agarose gel loaded with 1 Plus kb ladder and stained with SYBER Safe. 'M' indicates tracks loaded with the ladder, and 'Ctrl'- tracks loaded with PCR product amplified from WT LUHMES genomic DNA. PCR products derived from the isolated clones are labelled from 2.1 to 2.26. 'NTC' stands for non- template control. Expected PCR product size- 462bp. [B] Western blot analysis, assessing SORLA protein levels in two frameshift mutant clones (2.11 and 2.12) and one CRISPR WT line (2.9) compared to the protein levels of LUHMES WT line grown simultaneously. Membrane was probed for vinculin to verify equal loading. The expected protein sizes (SORLA- 330kDA and vinculin- 110kDA) are marked on the left of the image.

Table 6.2. A table listing all clones isolated following the second attempt to knock out SORL1 by targeting exon 31, the PCR and sequencing results, and the consequences of the mutations introduced.

Clone	PCR	Sequencing	Impact
2.1	Double band		
2.2	Single band	Sequencing failed	
2.3	Single band	Sequencing failed	
2.4	Three bands		
2.5	Single band	1bp insertion	Frameshift mutation
2.6	PCR failed		
2.7	Double band		
2.8	Double band		
2.9	Single band	CRIPSR WT	
2.10	Single band	2bp deletion	Frameshift mutation
2.11	Single band	5bp deletion	Frameshift mutation
2.12	Single band	4bp deletion	Frameshift mutation
2.13	Single band	4bp deletion	Frameshift mutation
2.14	Double band		
2.15	Single band	1bp insertion	Frameshift mutation
2.16	Dead		
2.17	Single band	1bp insertion	Frameshift mutation
2.18	Double band		
2.19	PCR failed		
2.20	Single band	1bp insertion	Frameshift mutation
2.21	Single band	15bp deletion	Inframe deletion
2.22	Single band	CRISPR WT	
2.23	Dead		
2.24	Dead		
2.25	Dead		
2.26	Double band		

6.3. Introducing the Gly508Ser mutation in LUHMES cells using CRISPR/Cas9 genome editing

Following the two unsuccessful attempts of generating *SORL1* KO clones using a gRNA targeting exon 31, I decided to move on to using CRISPR/Cas9 genome editing in an attempt to introduce the G508S mutation in *SORL1*. G508S is a rare coding *SORL1* variant, identified by Verheijen *et al.* (2016)[77] that is associated with autosomal dominant EOAD. It is found in individuals with AD with family history of AD, but not in age- and origin-matched controls[77]. The variant constitutes a point mutation (g.70501G>A), affecting an amino acid that is highly conserved across species (Figure 6.3.A). Importantly, it is located within close proximity to the binding site for Aβ[178]. The G508S variant has been predicted to be damaging by PolyPhen2, SIFT and MutationTaster[77]. Although it was not mentioned in the paper, I noted that the variant lies at the last nucleotide of *SORL1* exon 10 (Figure 6.3.A) and thus might also affect splicing. Therefore, it can be hypothesised that the G508S mutation might affect SORLA functions possibly by disrupting a splicing site and/or affecting its interaction with Aβ and thus its role in Aβ lysosomal targeting.

In addition, I also expected to obtain clones with frameshift mutations in exon 10 as by-products of introducing the G508S mutation. Any unrepaired CRISPR/Cas9-generated indels at this site might have altered splicing, leading to exon skipping or intron retention. Such events could have shifted the open reading frame, ultimately introducing premature stop codons and leading to activation of NMD pathways, and thus knocked out the full-length *SORL1* protein-coding transcript.

6.3.1. gRNA and ssODN design

As mentioned in chapter 4, gRNA is required for the targeting of the Cas9 endonuclease to the area of interest. Once delivered to the correct site, the endonuclease will create a DNA DSB, which will be repaired either by NHEJ or HDR mechanism, with the latter requiring the presence of a repair template. Exogenously introduced repair templates, mostly in the form of single stranded DNA oligonucleotides (ssODNs), have been utilised for introducing single-nucleotide mutations in order to investigate their functional consequences *in vitro*. This strategy has already been successfully implemented in LUHMES cells for introducing

missense mutations in the *MECP2* and *EEF1A2*- genes, associated with Rett syndrome and neurodevelopmental delay and intellectual disability, respectively[135].

6.3.1.1. Designing a suitable gRNA for introducing the G508S mutation into the LUHMES neuronal cell line

As before, gRNA was designed using the MIT Target Finder (http://crispr.mit.edu) and DNA2.0 (https://www.dna20.com/eCommerce/cas9/input) online CRISPR tools, and ligated into the pSpCa9(BB)-2A-GFP plasmid. However, this time, the gRNA was selected not only based on its predicted on- and off-target activity, but also considering the distance between the predicted cut site and the site of the intended missense mutation. Cut-to-mutation distance has been shown to play an important role in the efficiency of mutation incorporation. Both *in vivo* and *in vitro* studies have demonstrated that the frequency of the intended substitution declines with increasing the distance from the gRNA cut site[179,180]. Moreover, a paper published by Paquet *et al.* (2016)[180] suggests that this distance dependence could be manipulated for achieving a desired ratio of homozygous to heterozygous to wild type genotypes (Figure 6.3.B). According to this study, a cut-to-mutation distance in the range of 10bp to 20bp results in the highest number of heterozygous mutants and a good proportion of CRISPR WT clones. This is particularly important since the majority of the clinically relevant single-nucleotide substitutions, including the EOAD-associated G508S mutation, are heterozygous. Thus, given the results described by Paquet *et al.* (2016)[180] and in order to maximise the chances of obtaining a good proportion of heterozygote mutant, I selected a gRNA for which the distance between the Cas9 cut site and G508S mutation was 11bp (Figure 6.3.D).

6.3.1.2. Designing a suitable repair template for introducing the G508S mutation in the LUHMES neuronal cell line

The incorporation of a repair template, harbouring a mutation of interest, requires the activation of the HDR pathway, a process which, unlike the NHEJ, occurs at a considerably lower frequency in post-mitotic cells. A recently described method for increasing the efficiency of this process included designing an asymmetrical ssODN of a specific length. According to the study published by Richardson *et al.* (2016)[181], the Ca9 protein remains bound to the target DNA region for several hours. However, the 3' end of the PAM distal, non-target strand gets released much earlier than the other three strands, resulting in a 30bp free DNA stretch. Based on this observation,

Richardson *et al.*[181] showed that designing a template that would first anneal to the free, non-target strand could dramatically increase the frequency of HDR events (Figure 6.3.C). Consistent with previous studies[182,183], they found ssODNs complementary to the non-target strand much more efficient than those complementary to the target strand. In addition, Richardson *et al.*[181] were the first ones to suggest that asymmetric DNA donor might be more efficient in promoting HDR events, when compared to the commonly used symmetric around the break point ssODNs[182,183]. Bearing this in mind, I designed an asymmetric template, complementary to the non-target strand and overlapping the PAM-distal side with 36bp, and the PAM-proximal side with 91bp (Figure 6.3.D).

Another problem frequently encountered in CRISPR-mediated knock-in experiments is the presence of unwanted indels alongside the intended mutation, probably as a result of high Cas9 nuclease activity during multiple rounds of editing. Thus, in order to circumvent this problem, I introduced a silent point mutation at the PAM site as this has been previously demonstrated to be efficient in preventing multiple rounds of editing (Figure 6.3.D). The PAM site I mutated corresponds to the amino acid proline, which is encoded by four synonymous codons. Each of these codons are used at different frequencies, and substituting from frequent to rare codon has been shown to alter translation efficiency[184]. To account for this, I introduced a silent mutation which changed the codon to one with a similar usage frequency.

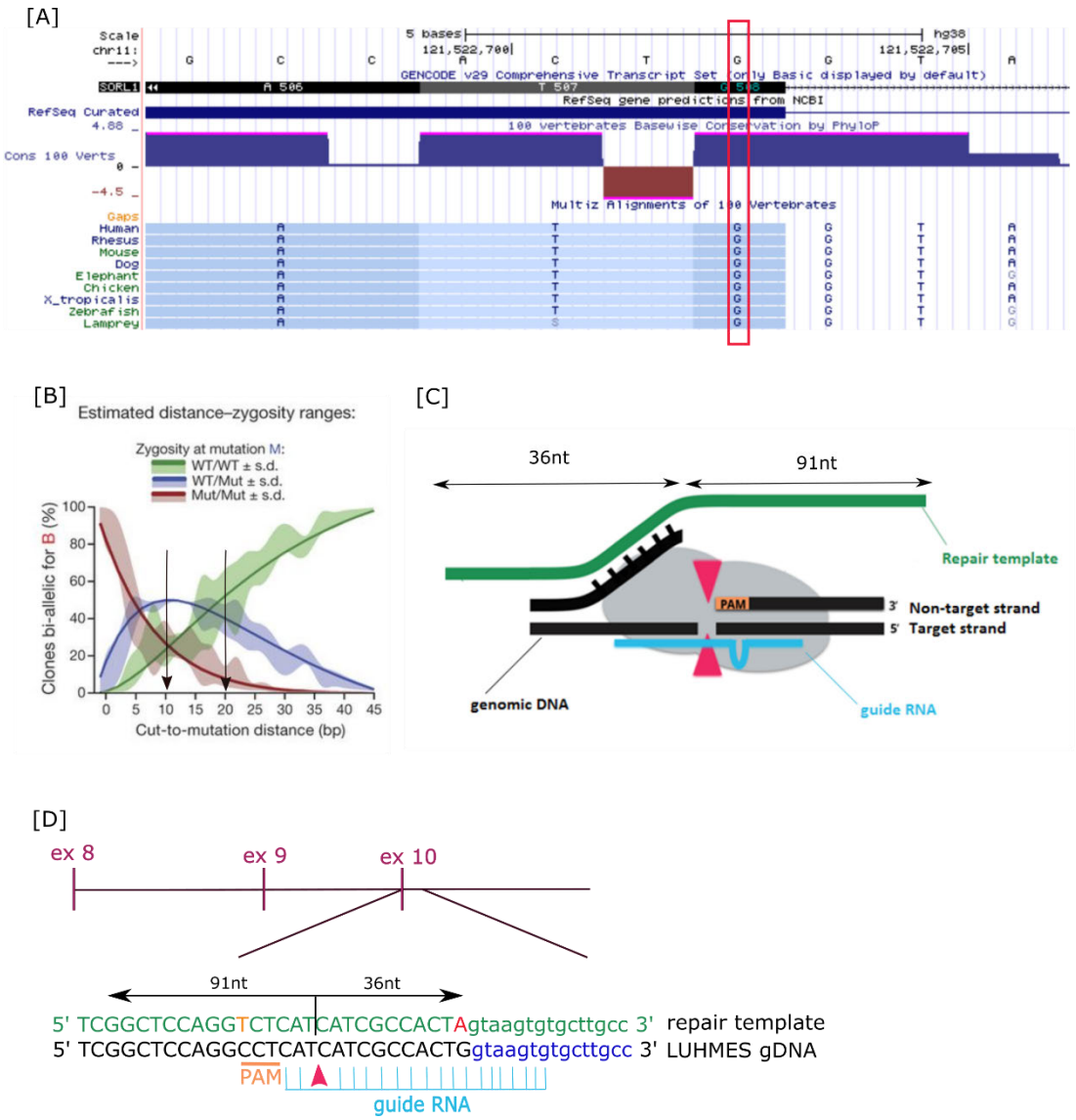

Figure 6.3. Summary of the experimental design of introducing the G508S mutation in SORL1 using CRISPR/Cas9 genome editing. [A] An image adopted after the UCSC genome browser (https://genome.ucsc.edu/) and showing the position of the affected by the G508S missense mutation nucleotide (marked with a red rectangular shape). First track shows the scale, the nucleotides and their co-ordinates in chromosome 11. Second track shows the affected by the mutation transcript, and the third track- the reference sequence corresponding to it. Tracks below show the conservation of the affected nucleotide amongst species. [B] A figure adopted after Paquet et al. (2016)[180], showing the relationship between the cut-to-mutation distance (x-axis) and the proportion of heterozygous for the mutation of interest clones (y-axis). Black arrows indicate the range of cut-to-mutation distance (bp) for obtaining an optimal number of heterozygous mutant and a good proportion of CRISPR WT clones. [C] A figure adopted after Richardson et al. (2016)[181], depicting the interaction of the asymmetric ssODN with the first DNA strand released after the Cas9-mediated DNA cleavage. [D] A schematic representation

of the designed ssODN for introducing the G508S mutation. The repair template (shown in green) was designed to be complementary to the non-target strand (shown in black with blue letters indicating an intronic region). The PAM site and the introduced silent mutation are shown in yellow, the point mutation in red, and the gRNA sequence in bright blue. The pink arrow indicates the predicted Cas9 cut site, with the designed ssODN being asymmetric relative to it.

6.3.2. Introducing the G508S mutation in LUHMES cells

The pSpCas9 (BB)-2A-GFP plasmid, containing the cloned gRNA, was transfected into proliferating LUHMES cells along with the G508S repair template. Forty-eight hours post-transfection, the top 3% of GFP-positive cells were FAC sorted into 96-well plates. Ten days after sorting, single cell colonies were identified. One third of each colony was kept for genotyping, and the rest of the cells were split into two 48-well plates for further expansion. Once 85% confluency was reached, the cell lines from one of the 48-well plates were frozen down, while the other plate was further expanded to 6-well plates before freezing.

6.3.2.1. First attempts to introduce the G508S mutation in LUHMES cells

The first attempt to introduce the G508S mutation in proliferating LUHMES cells had a transfection efficiency of 5.84% (measured as the percentage GFP-positive cells identified by FACS) and resulted in the formation of 30 single cell colonies. As described above, a proportion of cells were kept for genotyping. DNA was extracted by boiling the cells in Tris buffer and PCR was performed using primers flanking the G508S mutation site. High fidelity DNA polymerase was used for genotyping the identified cell lines in order to ensure that any mismatches identified by sequencing would be the result of the CRISPR process rather than error during PCR amplification.

From the 30 colonies identified, eight showed two bands on the gel (and were discarded as this suggested the presence of a large indel mutation), and two were discarded as they had three bands. Fifteen clones showed single bands when run on the gel and thus were subject to sequence analysis. The DNA isolated from four clones, however, failed to amplify (Figure 6.4.A). Of the 15 sequences obtained, only nine matched to the amplified area, and unfortunately none of them carried the intended G>A substitution.

Successful incorporation of the G508S mutation results in the formation of a SpeI restriction site. In order to quickly check whether any of the clones that failed to produce clean sequence traces (Table 6.3), carried the G508S mutation, I performed a restriction enzyme digest using the SpeI enzyme. The digests were run on a gel, and none of the samples showed evidence of incorporation of the G508S mutation (Figure 6.4.B). The results from the genotyping are summarised in Table 6.3.

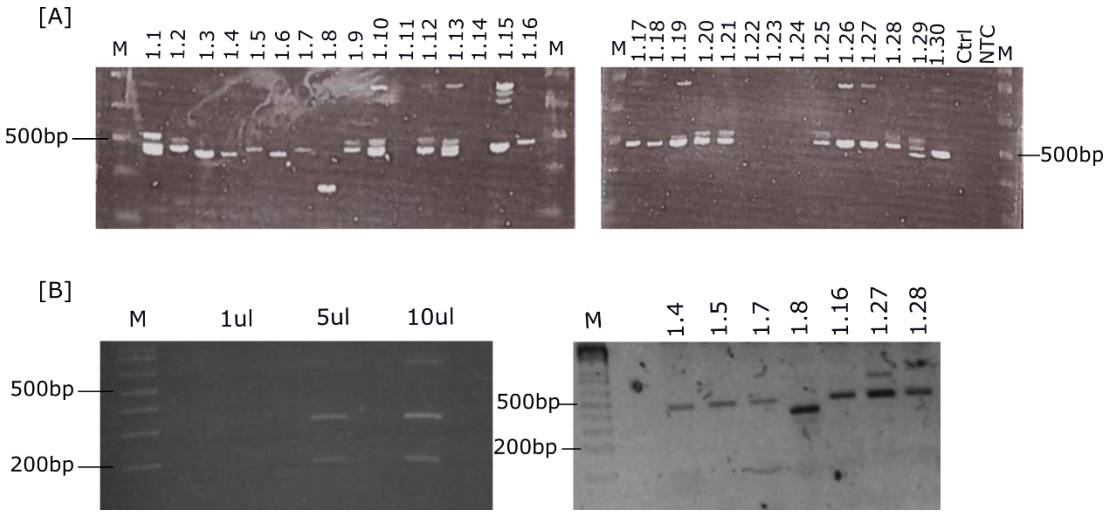

Figure 6.4. First attempt of introducing the G508S mutation in SORL1 using CRISPR/Cas9 genome editing. [A] Genotyping of the isolated clones, following FAC sorting. PCR with primers, flanking SORL1 exon 10, run on 2% LMP gel loaded with 1 Plus kb ladder and stained with SYBER Safe. 'M' indicates tracks loaded with the ladder. Due to pipetting error the 'Ctrl' track, corresponding to PCR-amplified WT LUHMES genomic DNA, was empty. PCR products of DNA derived from the isolated clones are labelled from 1.1 to 1.30. 'NTC' stands for non-template control. Expected product size 553bp. [B] Restriction enzyme digest on the clones, whose PCR products failed to give readable sequence traces. 1.8% LMP gels loaded with 1 PLus kb ladder (indicated as 'M') and stained with SYBER Safe. Left gel represents optimisation of the SpeI digest using a control PCR. Digesting 1µl of the control PCR did not give any visible products on the gel, using 5µl of the PCR reaction resulted in complete digest, while 10µl- incomplete. Subsequently, 5µl of PCR product for each clone were digested.

Table 6.3. A table listing the genotyping results for all clones isolated following the first attempt of introducing the G508S mutation in SORL1, the PCR and sequencing results and the consequences of the mutations introduced.

Clone	PCR	Sequencing	Impact
1.1	Three bands	Not sequenced, discarded	
1.2	Double band		
1.3	Single band	18bp deletion	Inframe deletion
1.4	Single band	Sequencing failed	SpeI digest did not identify presence of the G508S mutation
1.5	Single band	Sequencing failed	SpeI digest did not identify presence of the G508S mutation
1.6	Single band	17bp deletion	Frameshift mutant, high effect
1.7	Single band	Sequencing failed	SpeI digest did not identify presence of the G508S mutation
1.8	Single band	Sequencing failed	SpeI digest did not identify presence of the G508S mutation
1.9	Double band		
1.10	Double band		
1.11	PCR failed		
1.12	Double band		
1.13	Three bands	Not sequenced, discarded	
1.14	PCR failed		
1.15	Single band	16bp deletion	Inframe deletion
1.16	Single band	Sequencing failed	SpeI digest did not identify presence of the G508S mutation
1.17	Single band	1bp insertion	Frameshift mutation
1.18	Single band	Multiple deletions	
1.19	Single band	3bp deletion	Inframe deletion
1.20	Double band		
1.21	Double band		
1.22	PCR failed		

1.23	PCR failed		
Clone	**PCR**	**Sequencing**	**Impact**
1.24	PCR failed		
1.25	Double band		
1.26	Single band	1bp insertion	Frameshift mutation
1.27	Single band	Sequencing failed	SpeI digest did not identify presence of the G508S mutation
1.28	Single band	Sequencing failed	SpeI digest did not identify presence of the G508S mutation
1.29	Double band		
1.30	Single band	11bp deletion	Frameshift mutation

Since most of the lines from the first round of transfections were edited, but none of them carried the G508S mutation, I concluded that the gRNA was efficient in cutting, but perhaps the amount of the repair template used was insufficient. To test this, I carried out a second transfection, using 100uM instead of 10uM of the repair template. This time, the transfection efficiency dropped to 2.62% and only 19 single cell colonies were identified. To improve the quality of the PCR and subsequently to obtain cleaner sequencing results, I isolated DNA from the identified cell colonies using the Qiagen DNeasy kit rather than by boiling the cells. This time, most samples (apart from clones 2.1 and 2.2) amplified successfully. Six clones showed three bands on the gel and were discarded, four clones had two bands and seven clones- one (Figure 6.5). All samples that showed single bands on the gel were subject to sequence analysis and sent for sequencing. Sequencing identified clone 2.8 as a homozygote for the silent PAM site mutation, and clone 2.4 as a CRISPR WT. However, none of the clones carried the G508S. The results from the genotyping are summarised in Table 6.4.

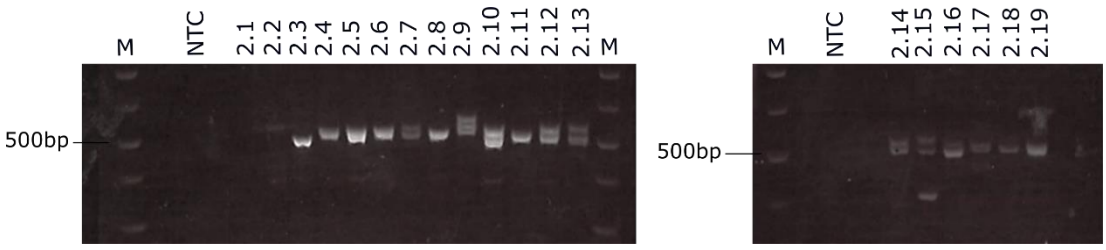

Figure 6.5. Genotyping of the isolated clones, following a second attempt to introduce the G508S mutation in SORL1 using CRISPR/Cas9 genome editing. 1.8% LMP gel loaded with 1 Plus kb ladder and stained with SYBER Safe. 'M' indicates tracks loaded with the ladder, and

'NTC'- non template control. PCR products of genomic DNA derived from the isolated clones are labelled from 2.1 to 2.19. Expected PCR product size- 553bp.

Table 6.4. A table listing the genotyping results for all clones after the second attempt of introducing the G508S mutation in SORL1, the PCR and sequencing results and the consequences of the mutation introduced.

Clone	PCR	Sequencing	Impact
2.1	PCR failed		
2.2	PCR failed		
2.3	Single band	26bp deletion	Inframe mutation
2.4	Single band	CRISPR WT	
2.5	Double band		
2.6	Double band		
2.7	Three bands	Not sequenced, discarded	
2.8	Single band	PAM site mutation (homozygous)	
2.9	Three bands	Not sequenced, discarded	
2.10	Three bands	Not sequenced discarded	
2.11	Single band	3bp deletion	Inframe mutation
2.12	Double band		
2.13	Three bands	Not sequenced, discarded	
2.14	Double band		
2.15	Three bands	Not sequenced, discarded	
2.16	Double band		
2.17	Three bands	Not sequenced, discarded	
2.18	Single band	11bp deletion	Frameshift mutation
2.19	Double band		

6.3.2.2. Successful incorporation of the G508S mutation in LUHMES cells

Since increasing the repair template amount coincided with a drop in the transfection efficiency, I decided to repeat the transfection using the same ssODN amount as in the first transfection, hoping to obtain more single cell colonies and in this way increase the chances of generating clonal lines carrying the G508S mutation. Even

though the transfection efficiency was low (1.22% of the FAC-sorted cells were GFP-positive), ten days post-transfection I was able to identify 95 colonies derived from single cells. As before, DNA was extracted, and PCR was run using high fidelity DNA polymerase, with primers flanking the mutation site.

Fourteen clonal lines were discarded as they showed three bands when the PCR reactions were run on a gel (Figure 6.6). Twenty-four clones were identified as heterozygotes as they had two bands of different sizes on the PCR gel (Figure 6.6), and another six clones showed double sequence traces around the Cas9 cut site. From the 50 clonal lines that showed a single band on the gel (Figure 6.6), 17 carried indels of various sizes and a couple showed mismatches around the mutation site. Five clonal lines remained WT after the CRISPR process and three were heterozygous for the G508S mutation (Figure 6.7). One clone carried the G508S mutation together with 2bp insertion on the same allele.

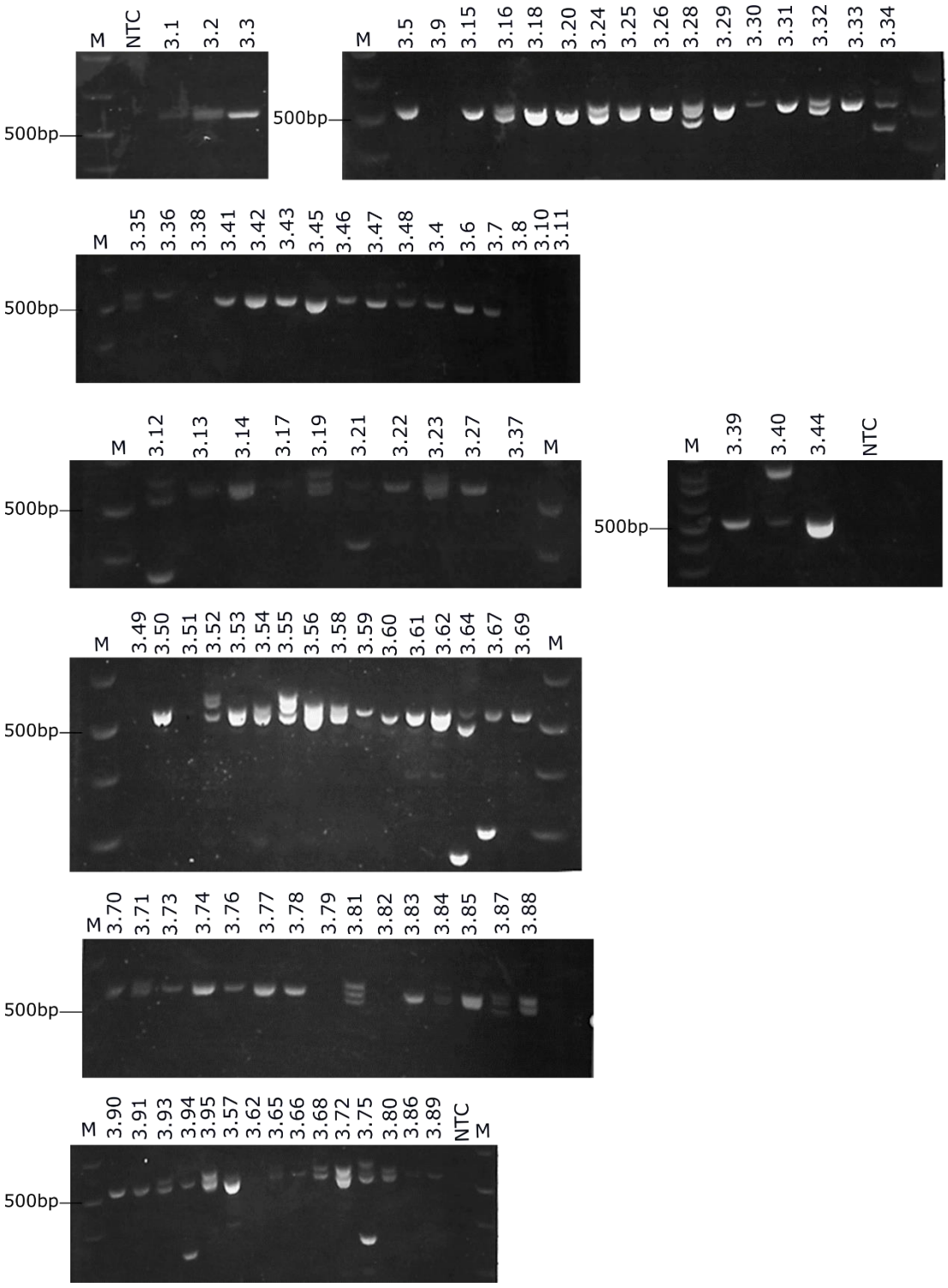

Figure 6.6. Results from the third attempt to introduce the G508S mutation in SORL1. 1.8% LMP gels loaded with 1 Plus kb ladder (indicated as 'M') and stained with SYBER Safe. PCR products of genomic DNA derived from the isolated clones are labelled from 3.1 to 3.95. 'NTC' stands for non-template control. Expected PCR product size- 553bp.

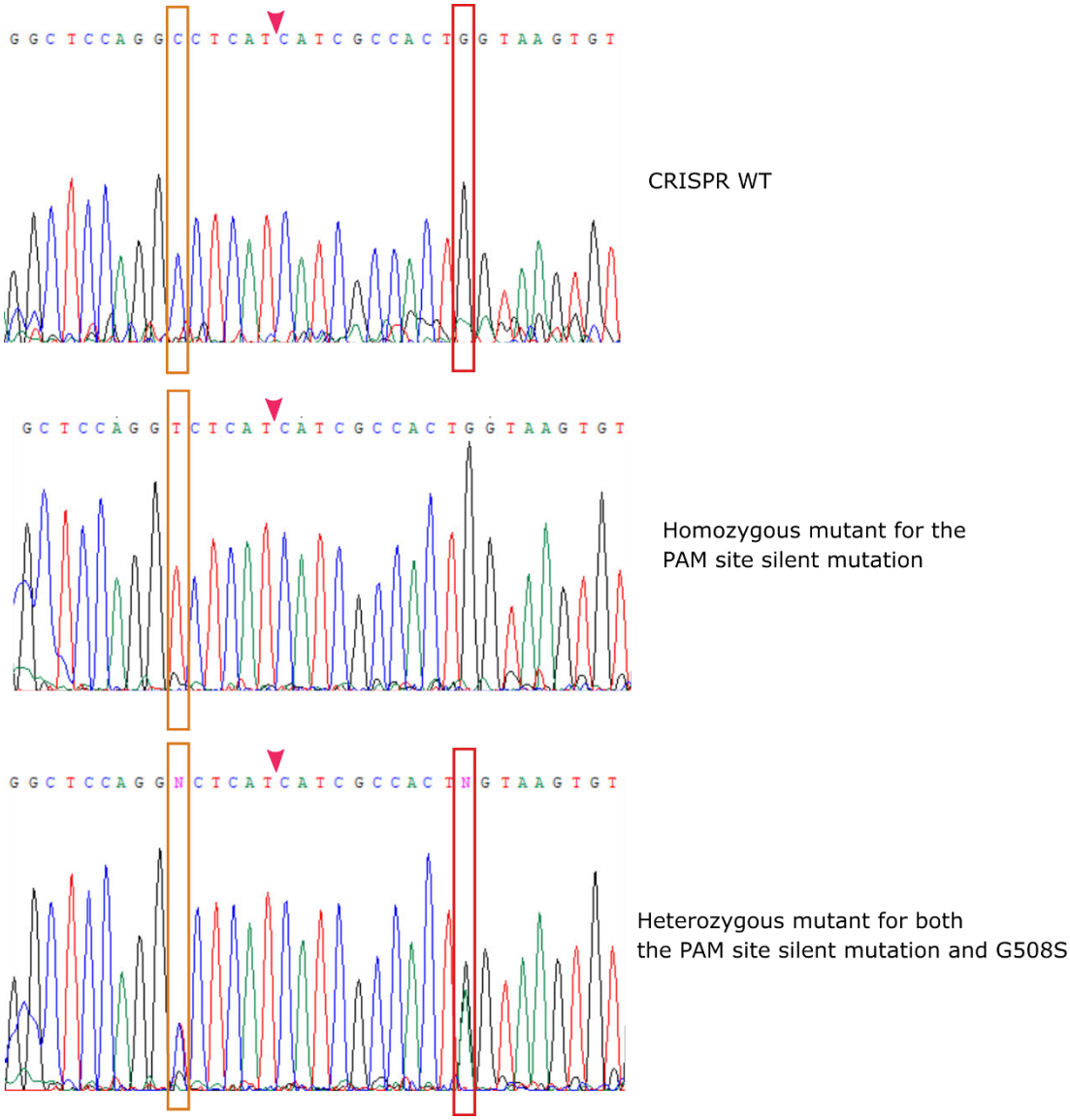

Figure 6.7. Representative sequencing traces. Yellow rectangles indicate the point of the introduced silent PAM site mutation, red- the site of the G508S point mutation, and red arrows- the predicted Cas9 cut site. Top trace corresponds to a CRISPR WT clone, middle trace corresponds to a CRISPR clone that was homozygous for the PAM site mutation, and bottom trace show the sequence of a CRISPR clone that was identified as heterozygous for both the PAM site and the G508S mutation.

In order to determine whether some of the clones that showed two alleles on the gel were in fact compound heterozygotes carrying the G508S mutation on one allele, I performed restriction enzyme digest with SpeI as described in 6.3.2.1. From all samples, only three clonal lines- 3.44, 3.58 and 3.71 were digested (Figure 6.8.).

Thus, it was concluded that they constitute compound heterozygotes, carrying the G508S mutation on one allele, and an indel on the other. The genotyping results for all clones are summarised in Table 6.5.

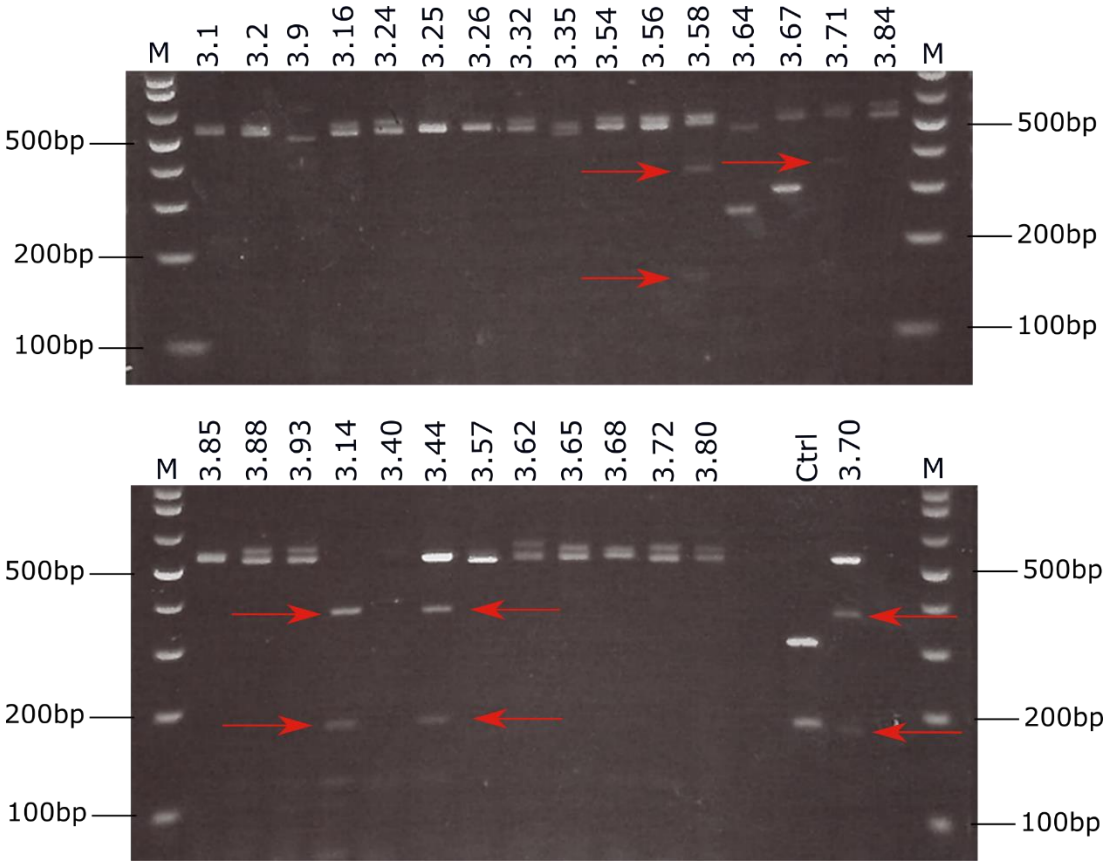

Figure 6.8. Identifying compound heterozygous mutants, carrying the G508S mutation using restriction enzyme digest. 1.8% LMP gels loaded with 1 Plus kb ladder (indicated as 'M') and stained with SYBER Safe. Clones that showed double bands on the genotyping gel (Figure 6.6.) were digested with the SpeI enzyme. Red arrows indicate the resulted in digestion products in the clones, which carried the G508S mutation. A control PCR (labelled as 'Ctrl') and the 3.70 clone, which was identified as G508S heterozygous mutant by sequencing, were used as positive controls. Expected PCR product size 553bp; expected digest products size: 389bp and 164bp.

Table 6.5. A table listing the PCR and sequencing results from genotyping the isolated clones after the third attempt of introducing the G508S mutation in SORL1. The type and the consequences of the introduced mutations are also included.

Clone	PCR	Sequencing	Impact
3.1	Double band		
3.2	Double band		
3.3	Single band	9bp deletion	Inframe deletion
3.4	Single band	Multiple mismatches	
3.5	Single band	PAM site mutation (homozygous)	
3.6	Single band	PAM site mutation (homozygous)	
3.7	Single band	PAM site mutation (homozygous)	
3.8	PCR failed		
3.9	Double band		
3.10	PCR failed		
3.11	Three bands	Not sequenced, discarded	
3.12	Three bands	Not sequenced, discarded	
3.13	Single band	Matches *SORL1* exon 11	
3.14	Single band	Multiple mismatches	
3.15	Single band	3bp deletion	Inframe mutation
3.16	Double band		
3.17	Single band	Double sequencing trace around the cut site	
3.18	Single band	Double sequencing trace around the cut site	
3.19	Three bands	Not sequenced, discarded	
3.20	Single band	PAM site mutation (homozygous)	
3.21	Three bands	Not sequenced, discarded	
3.22	Single band	'Unclean' sequence trace	
3.23	Three bands	Not sequenced, discarded	
3.24	Double band		
3.25	Double band		
3.26	Double band		
3.27	Single band	3bp deletion	Inframe deletion
3.28	Three bands	Not sequenced, discarded	

Clone	PCR	Sequencing	Impact
3.29	Single band	3bp deletion	Inframe deletion
3.30	Single band	3bp deletion	Inframe deletion
3.31	Single band	PAM site mutation + G508S mutation (heterozygous)	
3.32	Double band		
3.33	Single band	CRISPR WT	
3.34	Three bands	Not sequenced, discarded	
3.35	Double band		
3.36	Single band	Sequencing failed	
3.37	PCR failed		
3.38	PCR failed		
3.39	Single band	3bp deletion	Inframe deletion
3.40	Double band		
3.41	Single band	3bp deletion	Inframe deletion
3.42	Single band	3bp deletion	Inframe deletion
3.43	Single band	PAM site mutation (homozygous)	
3.44	Double band	G508S mutation (compound heterozygous)	
3.45	Single band	Multiple deletions	
3.46	Single band	PAM site mutation (homozygous)	
3.47	Single band	6bp deletion	Inframe deletion
3.48	Single band	3bp deletion	Inframe deletion
3.49	PCR failed		
3.50	Single band	1bp mismatch	
3.51	Single band	Sequencing failed	
3.52	Three bands	Not sequenced, discarded	
3.53	Single band	Double sequence trace	
3.54	Double band		
3.55	Three bands	Not sequenced, discarded	
3.56	Double band		
3.57	Three bands	Not sequenced, discarded	
3.58	Single band	G508S mutation (compound heterozygous)	

Clone	PCR	Sequencing	Impact
3.59	Single band	G508S mutation + 2bp insertion (homozygous)	
3.60	Single band	11bp deletion	Frameshift mutation
3.61	Single band	PAM site mutation (homozygous)	
3.62	PCR failed		
3.63	Single band	Double sequence trace	
3.64	Three bands	Not sequenced, discarded	
3.65	Double band		
3.66	Single band	3bp deletion	Inframe deletion
3.67	Double band		
3.68	Double band		
3.69	Single band	PAM site mutation (homozygous)	
3.70	Single band	PAM site mutation (homozygous) + G508S mutation (heterozygous)	
3.71	Single band	G508S mutation (compound heterozygous)	
3.72	Double band		
3.73	Single band	Double sequence trace	
3.74	Single band	CRISRP WT	
3.75	Three bands	Not sequenced, discarded	
3.76	Single band	PAM site mutation (heterozygous)	
3.77	Single band	3bp deletion	Inframe deletion
3.78	Single band	3bp deletion	Inframe deletion
3.79	PCR failed		
3.80	Double band		
3.81	Three bands	Not sequenced, discarded	
3.82	PCR failed		
3.83	Single band	3bp deletion	Inframe deletion
3.84	Double band		
3.85	Double band		
3.86	Single band	4bp insertion	Frameshift mutation

Clone	PCR	Sequencing	Impact
3.87	Three bands	Not sequenced, discarded	
3.88	Double band		
3.89	Single band	Sequencing failed	
3.90	Single band	CRISPR WT	
3.91	Single band	CRISPR WT	
3.92	Single band	1bp insertion	Frameshift mutation
3.93	Double band		
3.94	Double band		
3.95	Double band		

6.4. Assessing the effect of the G508S mutation on *SORL1* expression and Aβ production

Having knocked in the G508S mutation into LUHMES cells, I wanted to investigate its functional consequences in terms of *SORL1* expression and of its impact on APP processing. In order to do this, I defrosted all clonal lines that were heterozygotes for the G508S mutation (clones 3.14, 3.31, 3.61 and 3.70), as well as the compound heterozygotes (3.44, 3.58 and 3.71) and the CRISPR WTs (3.33, 3.74, 3.90, 3.91). Unfortunately, in the process of defrosting, all knock-in lines but one, 3.70, died. The same happened with the CRISPR WTs, where only clone 3.90 survived, while all the compound heterozygotes survived. I differentiated clones 3.70, 3.90, 3.44, 3.58 and 3.71 into neurons and performed all downstream experiments described below 14 days post-differentiation.

6.4.1. Assessing the effect of the G508S mutation on *SORL1* expression levels

6.4.1.1. Assessing the effect of the introduced G508S mutation on splicing

As mentioned in the introduction, the G508S mutation constitutes a G>A substitution at the last nucleotide of *SORL1* exon 10. This nucleotide is part of a 5' splice site

(CAG/GUAAGU sequence). Substitutions in this canonical sequence may alter splicing by affecting the interaction between pre-mRNA and proteins involved in intron removal[185].

In order to investigate whether the G508S mutation affects *SORL1* splicing, I extracted RNA from WT LUHMES neurons (day 14), as well clones 3.90 (CRISPR WT) and 3.70 (G508S heterozygote), and reverse transcribed it to cDNA. Then, I performed an RT-PCR, using primers, located in *SORL1* exons 9 and 12. These primers would be expected to give a product size of 395bp. However, if the G508S mutation led to the intron following exon 10 being retained, the amplified product would be predicted to be 607bp. As can be seen on Figure 6.9.A, the G508S mutation, does not affect *SORL1* splicing, as PCR amplification produced a single band of the same size as that observed in WT LUHMES and brain cDNA at the predicted cDNA-derived size of 395bp.

6.4.1.2. Assessing *SORL1* mRNA levels in the G508S heterozygous and compound heterozygous clones

As mentioned above, coding variants in *SORL1* implicated in AD have been shown to reduce *SORL1* transcript levels. This has been attributed to either activation of NMD pathways or alteration of translation efficiency[80,77,177]. Therefore, despite the fact that the G508S mutation did not alter splicing, I decided to test whether it exerted an effect on *SORL1* mRNA and/or protein levels.

First, I assessed how the introduced G508S mutation affected *SORL1* mRNA levels. I grew and differentiated clones 3.90 (CRISPR WT), 3.70 (G508S heterozygote mutant), 3.44, 3.58 and 3.71 (G508S compound heterozygote mutants), alongside LUHMES WT lines of a similar passage number. RNA was extracted from day 14 neurons and reverse transcribed to cDNA. All clones were grown and differentiated three independent times. Some of the clones died before reaching day 14. Thus, for the G508S compound heterozygous clones 3.44 and 3.58 I was only able to collect RNA pellets from two, and for clone 3.71- from one independent growth. Subsequently, I performed qRT-PCR using two TaqMan probes, located near the 5' (*SORL1* exon 2-3) and the 3' end (*SORL1* exon 46-47) of the full- length *SORL1* transcript. The geometric mean of two reference genes (*UBE4* and *TBP*) selected from a panel of eight, as the most stable in this sample set by the GeNorm software, was used to normalise *SORL1* expression levels. All assessed clones, including the

CRISPR WT, showed significant reduction in *SORL1* mRNA levels compared to simultaneously grown LUHMES WT lines using the TaqMan probe detecting all *SORL1* transcripts (amplification of *SORL1* exon 46-47) (Figure 6.9.B, right). Meanwhile, clones 3.90 (CRISPR WT) and 3.70 (G508S heterozygote mutant) were the only ones showing a statistically significant decrease in *SORL1* expression when assessed with the probe detecting only the full-length *SORL1* transcripts (*SORL1* exon 2-3) (Figure 6.9.B, left).

I hypothesised that the approximately 50% reduction in *SORL1* expression observed in the CRISPR WT clone 3.90 could be due to haploinsufficiency caused by a large deletion within the targeted locus. A recent paper reported the presence of large deletions and genomic re-arrangements around CRISPR/Cas9-targeted sites that had not been predicted as off-targets by any of the available CRISPR/Cas9 online tools [186]. If occurring in one allele, such extensive genomic damage could lead to allele-specific loss of expression. Therefore, in order to investigate whether both *SORL1* alleles were present in the CRISPR WT clone 3.90, I designed a PCR-based screening assay. Using the UCSC genome browser (https://genome.ucsc.edu/), I identified four SNPs within different *SORL1* exons where the minor allele frequency was nearly 50%. In order to determine whether the LUHMES cell line was heterozygous for any of these four SNPs, I designed primers flanking them and ran RT-PCR using LUHMES WT cDNA as a template. Sequence analysis revealed that LUHMES were heterozygote for one of the four SNPs, which is located in *SORL1* exon 6. Subsequently, I used this primer pair to screen cDNA derived from the CRISPR WT (clone 3.90). Unlike the LUHMES WT line, clone 3.90 was homozygous for this SNP, indicating that this clone had lost one *SORL1* allele as a result of the CRISPR process. Thus, it was concluded that clone 3.90 did not constitute a CRISPR WT (Figure 6.9.C).

6.4.1.3. No difference in the SORLA protein levels was observed in the G508S heterozygote and compound heterozygote mutant lines

Since incorporation of the G508S mutation led to decreased *SORL1* mRNA levels in both heterozygote and compound heterozygote clones, I wanted to further investigate whether such reduction would also be observed at the protein level. I grew and differentiated the G508S clones (i.e. 3.90, 3.70, 3.44, 3.58 and 3.71), together with LUHMES WT controls of a similar passage number three independent times. Some of the clones (3.70, 3.44 and 3.71) died during one of the three growths. Following 14

days of differentiation, I collected protein lysates from these lines and performed western blot analysis. No significant difference in the SORLA protein levels was observed between either the G508S heterozygote clone, 3.70, or any of the three compound heterozygote clones (3.44, 3.58 and 3.71) when compared to WT LUHMES neurons. Additionally, clone 3.90, also showed SORLA levels similar to the WT controls (Figure 6.9.D).

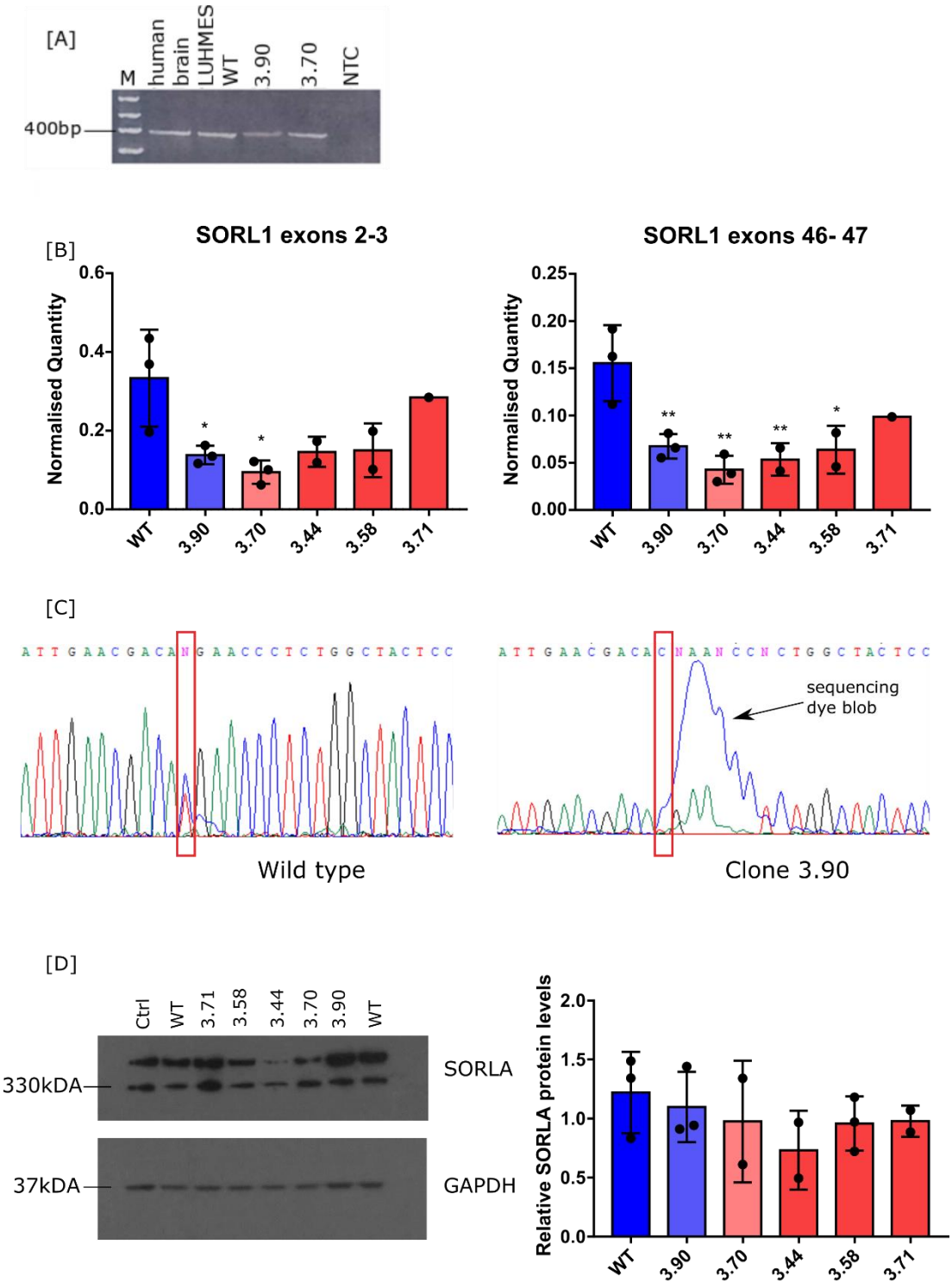

Figure 6.9. Assessing SORL1 expression in the G508S heterozygous and compound heterozygous clones. [A] RT-PCR assessing the effect of the G508S mutation on splicing. 1.7% agarose loaded with 1 Plus kb ladder (indicated as 'M') and stained with SYBER Safe. Lanes loaded with amplified cDNA products derived from human brain, LUHMES WT cells, CRISPR WT (3.90) and G508S heterozygous mutant (3.70) clones. Expected product size-397bp. [B] qRT-PCR assessing the SORL1 mRNA levels in the isolated clones. The x-axis

shows the clones, while the y-axis- SORL1 mRNA levels normalised to the geometric mean of UBE4 and TBP. WT LUHMES lines are shown in blue, clone 3.90 in bright blue, clone 3.70 (G508S heterozygous mutant) in pink and clones 3.44, 3.58 and 3.71 (compound heterozygotes) in red. Graphs represent average of three growths with each dot indicating an independent growth [C] Sequencing traces showing the rs12364988 SNP located in SORL1 exon 6 (marked with a red rectangular shape. 'N' at this position stands for any nucleotide and indicates that LUHMES are heterozygous for this SNPs. Meanwhile, clone 3.90 was identified as a homozygous for this SNP. [D] Western blot analysis, assessing SORLA protein levels in clone 3.90, the identified G508S heterozygote (3.70) and the three compound heterozygous mutants (3.44, 3.58 and 3.71), as well as in two LUHMES WT lines grown simultaneously. A representative western blot image is shown on the left, where a protein sample from IPSC-derived neurons was included as a positive control (labelled as 'Ctrl'). Membrane was probed for GAPDH to verify equal loading. The expected protein sizes (SORLA- 330kDA and GAPDH-37kDA) are marked on the left of the image. Quantification of the relative SORLA protein levels, normalised to GAPDH, in lysates collected from three independent growths (each dot on the graph represents an independent growth) is shown on the right. Error bars represent mean ± SD of three experimental replicates grown independently.

6.4.1.4. Identifying *SORL1* knockdown clones amongst the clones generated during the third attempt of introducing G508S

As summarised in table 6.5, a large proportion of the clonal lines, generated during the process of introducing the G508S mutation, appeared from the PCR analysis to carry an indel in at least in one of their alleles. In order to investigate whether some of these constituted *SORL1* KOs for the full-length *SORL1* protein coding transcript, Divya Pandya (an honours student working under my supervision) thawed four clones that were identified as homozygous frameshift mutants, as well as twelve lines that carried an indel in at least one allele as identified by PCR. Only ten clones survived thawing. These lines were expanded, RNA was isolated from proliferating cells and qRT-PCR was performed to assess *SORL1* mRNA levels. The cut site was located in exon 10, which is present only in the full-length *SORL1* transcript. Given this, both probes, amplifying the 3' (*SORL1* exon 2-3) and the 5' end (*SORL1* exon 46-47) of the full-length *SORL1* transcript were used to assess *SORL1* expression levels. *SORL1* expression was detected in all clonal lines with both probes (Figure 6.10.A). Five clones showed elevated *SORL1* mRNA levels (Figure 6.10.A). Meanwhile, clones 3.1, 3.16 and 3.40 displayed lower *SORL1* expression when compared to a

WT line of a similar passage number and grown at the same time (Figure 6.10.A). This decrease was detected using both probes (Figure 6.10.A). Subsequently, clones 3.1, 3.16 and 3.40 which had the lowest mRNA levels were differentiated, together with one of the clones that showed increased *SORL1* mRNA levels- 3.26, and two LUHMES WT lines of a similar passage number. Western blot analysis of protein lysates collected from these clones 14 days after differentiation identified clones 3.1 and 3.16 as potential *SORL1* knockdowns (Figure 6.10.B).

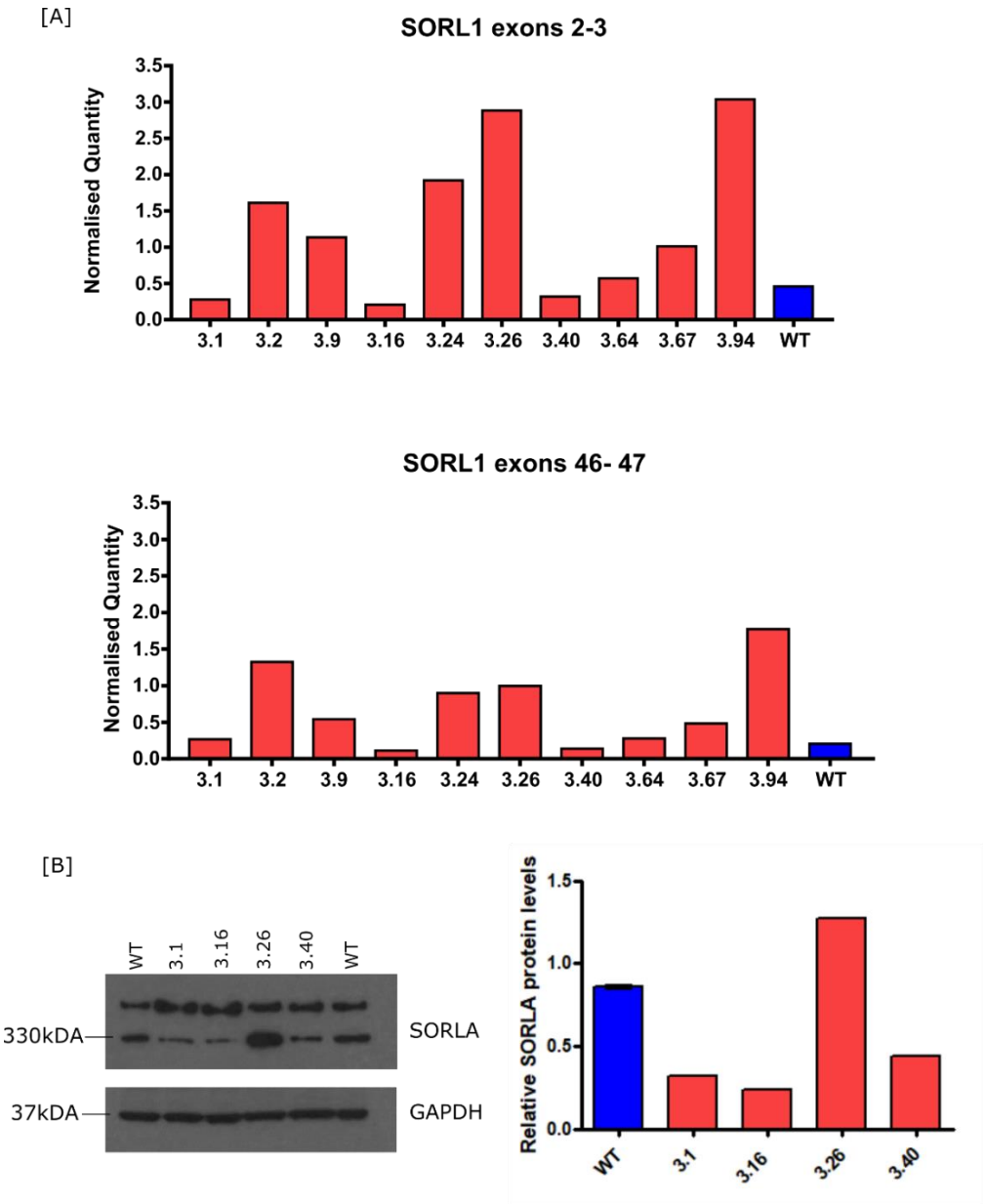

Figure 6.10. Identifying SORL1 knockdown clones. [A] qRT-PCR assessing the SORL1 mRNA levels in RNA from proliferating cell for the selected mutant clones. The x-axis shows the

clones, while the y-axis- SORL1 mRNA levels normalised to the geometric mean of PRDM5 and ENOX3. LUHMES WT is shown in blue and mutant clones in red. [B] Western blot analysis, assessing SORLA protein levels in LUHMES neurons (day 14) of the clones that showed the lowest SORL1 mRNA levels- 3.1, 3.16 and 3.40, as well as clone 3.26 and two LUHMES WT lines grown simultaneously. Western blot image is shown on the left with the expected protein sizes (SORLA- 330kDA and GAPDH- 37kDA) are marked on the left of the image. GAPDH was used to verify equal loading. Quantification of the relative SORLA protein levels normalised to GAPDH in the assessed clones is shown on the right. WT is shown in blue and mutant clones in red. Graphs represent the results from a single experiment.

6.4.2. Assessing the effect of the Gly508Ser on APP processing

SORLA plays a role in APP processing[37] and Aβ metabolism[41]. Moreover, many rare variants in *SORL1,* found in EOAD and LOAD patients, have been associated with altered APP processing and Aβ accumulation[41,76,173]. Therefore, I decided to examine the effect of the introduced G508S mutation on Aβ production. I grew and differentiated the G508S heterozygote (3.70) and compound heterozygote (3.44, 3.58 and 3.71) clones, alongside the identified *SORL1* knockdown lines (3.1 and 3.16) and three WT controls of a similar passage number. Unfortunately, the G508S heterozygote clone (3.70) died before reaching day 14. At day 14 of differentiation, I collected media from these lines and performed $A\beta_{42}$ sandwich ELISA. However, all of the results obtained were below the assay detection limit of 15.6pg/ml (Table 6.6) and were thus considered invalid.

Table 6.6. A table summarising the concentration of extracellular Aβ42 (pg/ml) obtained for each of the tested WT and SORL1 mutant cell lines using ELISA. All samples measured were below the assay lowest detection limit of 15.6pg/ml.

Cell line	Extracellular $A\beta_{42}$ concentration (pg/ml)
WT1	2.86
WT2	2.50
WT3	2.80
Clone 3.1	3.16
Clone 3.16	3.35
Clone 3.44	2.86
Clone 3.58	2.74
Clone 3.71	3.16

6.5. Discussion

The aims of the work described in this chapter were to:

a) knock out the *SORL1* gene and introduce the rare EOAD-associated *SORL1* missense mutation- G508S, in LUHMES cells using the CRISPR/Cas9 genome engineering system;
b) assess the effect of the introduced G508S missense mutation on *SORL1* expression;
c) investigate the effect of both knocking out *SORL1* and introducing the G508S missense mutation on Aβ production.

First, I decided to knockout *SORL1* by targeting exon 31, common to all *SORL1* transcripts found in LUHMES. Both attempts of knocking out *SORL1* using this gRNA seemed to prove unsuccessful. However, the first transfection generated two clones (1.2 and 1.3) with *SORL1* mRNA levels much lower than a LUHMES WT line grown simultaneously that may be useful in future analysis. Moreover, the second round of transfections resulted in many clones (eight out of 26 identified) that carried frameshift mutations in the targeted area. Unfortunately, due to time limitations, I screened only two of these clones further. Therefore, it is possible that some of these clones constitute *SORL1* knockouts or at least knockdowns. However, since frameshift

mutations introduced close to the carboxy terminal are less likely to activate NMD pathways and thus to knock out the gene (as discussed in 4.4.), and due to time restraints, this was not investigated further.

Introducing the G508S EOAD-associated missense mutation in *SORL1* into LUHMES cells using CRISPR/Cas9 genome editing also proved to be challenging, with the first two rounds of transfections not generating a single G508S mutant clone. The third round, however, had a success rate for incorporation of the G508S mutation of around 3%, which is within the range of knock-in efficiency previously reported for LUHMES cells (2%-8%)[135]. In comparison, the rate of incorporating the PAM site silent mutation was much higher (11%), suggesting that shorter cut-to-mutation distance might have resulted in better repair efficiency (the distance between the PAM site mutation and the cut site was 6bp, while that between the G508S mutation and the cut site- 11bp). All clones that carried the PAM site mutation, except for one, were homozygous. This is in agreement with the results of Paquet *et al.* (2016)[180], according to which decreasing the cut-to-mutation distance results in a higher proportion of homozygous mutants compared to the number of heterozygous and CRISPR WT clones[180].

The third attempt of introducing the G508S mutation generated four G508S heterozygous mutants, three compound heterozygotes (carrying the G508S mutation on one *SORL1* allele and an indel on the other), and four CRISPR WT clones. However, only one G508S heterozygous mutant and one CRISPR WT clone were revived following a short-term storage at -80^{o}C. The high death rates following thawing could be explained either by incorrect freezing conditions or by the fact that around the time most of the clones were frozen down the freezer broke down.

As mentioned above, the G508S missense mutation constitutes a single nucleotide substitution at the end of *SORL1* exon 10, thus affecting a 5$^{'}$ splice region. Therefore, it was hypothesised that the missense mutation might affect *SORL1* splicing and thus expression and function. RT-PCR with primers flanking *SORL1* exon 10 did not show any changes in splicing. Meanwhile, qRT-PCR detected lower *SORL1* mRNA levels in both the G508S heterozygous (3.70) and compound heterozygous (3.44, 3.58 and 3.71) clones. Surprisingly, clone 3.90, initially identified as a CRISPR WT, showed approximately 50% reduction in *SORL1* mRNA levels compared to the WT levels. A SNP-based RT-PCR assay showed that this clone is homozygous for a SNP in this area, unlike the control LUHMES WT cell line. Since the SNP used in this assay is located in *SORL1* exon 6, some distance from exon 10, the loss of heterozygosity

suggested the presence of a large deletion on one *SORL1* allele as a result of the CRISPR editing process. Such events have been recently reported by others[186] and have raised the question of the specificity of the CRISPR/Cas9 system. Given these results, clone 3.90 was excluded from any further analysis. Meanwhile, the decrease in *SORL1* expression detected in the G508S heterozygote and compound heterozygote clones did not appear to translate into reduced protein levels.

The location of the G508S mutation is at a splice site, which led to the hypothesis that indels in this region could potentially alter splicing and knockout the full-length *SORL1* protein transcript by activating NMD pathways. Thus, all clones that carried indels (identified either by sequencing analysis or based on the PCR results) and survived the thawing process were further screened to determine SORLA protein levels. Western blot analysis detected lower levels of SORLA protein in two of the clones that showed reduced *SORL1* mRNA levels by qRT-PCR. These clones (3.1 and 3.16) were identified as *SORL1* knockdowns.

Many rare mutations in *SORL1* have been linked to altered levels of APP processing and Aβ production[41,76,173]. In addition, the G508S mutation is located within close proximity of the SORLA binding site for Aβ[178]. Therefore, I hypothesised that introducing this missense variant will alter the levels of extracellular Aβ. Unfortunately, I was unable to detect $A\beta_{42}$ in the media collected from the G508S mutant clones, as well as from the WT and *SORL1* knockdown lines grown in parallel. These results suggest that a more sensitive $A\beta_{42}$ assay would need to be used in the future if this experiment was to be repeated. Alternatively, both $A\beta_{40}$ and sAPPβ have been previously detected in LUHMES neurons (day 12)[134]. Thus, it would be interesting to see whether introducing the G508S mutation has an effect on their production.

Chapter 7 Discussion

Summary of thesis

SORCS2 and SORLA are members of the Vps10p-domain receptor family, also known as the sortilins, which comprises a family of five receptors[2]. These receptors are predominantly expressed in the CNS, where they facilitate the sorting and trafficking of a diverse range of neurotrophic factors and their precursor forms, as well as of various transmembrane receptors and synaptic proteins[2]. All members of the receptor family have been implicated in multiple physiological process related to neuronal survival, differentiation and synaptic plasticity. Amongst these, the role of the receptor family in the sorting and processing of APP, and thus in Aβ generation was pivotal for the work described in this thesis. As discussed in chapter 1, genetic variation within the receptor family has been associated with many brain conditions, and knockout mice for these genes have provided further evidence for their involvement in neurodegenerative and psychiatric disorders, including AD.

Increased DNA damage and compromised DNA repair have also been implicated in the aetiology of both psychiatric[129] and neurodegenerative conditions[169], AD in particular. However, the majority of studies report evidence for oxidative DNA damage occurring as a result of increased production of free radicals and reduced activity of relevant DNA repair pathways[129,169]. More recently, a specific type of DNA damage-DNA DSB formation, has been linked to both AD and physiological brain activity, i.e. learning and memory processes. Suberbielle *et al.* (2013)[107] demonstrated that in mice, naturally occurring behaviour, such as exploration of a novel environment, leads to the transient formation of DSBs in the dentate gyrus (an area of the brain important for learning and memory), which are repaired within 24 hours. Similarly, Madabhushi *et al.* (2015)[118] detected elevated levels of DNA DSBs in hippocampal lysates from WT mice subjected to a fear conditioning paradigm, as well as in WT mouse hippocampal primary neurons stimulated with NMDA. Moreover, they showed that the generation of these breaks depends on TopoIIβ activity and results in an increase in early-response genes expression[118].

Meanwhile, elevated levels of neuronal DSBs and defective DSB repair have been found in the dentate gyrus of hAPP mice (an AD mouse model). Further *in vivo* and *in vitro* experiments identified aberrant network activity and dysregulated NMDAR signalling as underlying mechanisms[107]. Subsequently, increased levels of DSBs

have been also detected in the hippocampus and the frontal cortex of MCI and AD patients compared to age-matched cognitively unimpaired controls[128].

The aim of my project was to investigate the link between two members of the Vps10p-domain receptor family- SORCS2 and SORLA, and DNA DSB formation, as a possible mechanism in neuropsychiatric illness. This aim stemmed from a few research findings. First, as briefly described above and more extensively in the previous chapters, both the Vps10-domain family, including SORCS2 and SORLA, and DNA DSB formation have been associated with neuropsychiatric conditions, as well as with processes related to learning and memory. Second, SORCS2 (evidence from *in vitro* studies), and particularly SORLA (*in vitro* and *in vivo*) play a role in APP processing and Aβ generation, and accumulation of the latter exacerbates DNA DSB formation *in vivo* and *in vitro*. Third, SorCS2 has been implicated in the trafficking of the NR2A-containing NMDARs, shown to be required for DNA DSB repair *in vitro*.

In this chapter I will discuss my results, their limitations, as well as possible future directions towards examining further the relationship between SORCS2 and SORLA, DNA DSB formation and neuropsychiatric disorders.

Aim 1: Examining the levels of DNA DSB formation in WT, *Sorcs2*$^{-/-}$ and *Sorl1*$^{-/-}$ mice, at baseline and following exploration of a novel environment

To begin with, I replicated previous findings that naturally occurring behaviour, such as exploration of a novel environment, triggers transient increase in DSB formation in the hippocampus of adult WT mice. To do this, I utilised mouse brain samples obtained from our collaborator Assoc. Prof Simon Glerup (Aarhus University, Denmark), whose group had performed the behavioural experiments described below. I detected a significant increase in the percentage of nuclei positive for γH2A.X in the dentate gyrus, as well as in the CA2 and CA3 region of the hippocampus of WT mice, exposed to a novel environment compared to WT mice that remained in their home cages. Moreover, when the mice were left to recover for 24 hours after the exploratory activity in their home cages, the percentage of γH2A.X-positive neurons in the examined brain areas dropped back to baseline levels. However, the novel finding in this study was the effect of knocking out *Sorcs2* on this process. *Sorcs2*$^{-/-}$

mice in the home cage group showed significantly higher levels of baseline DNA DSBs in the dentate gyrus, but not in CA2 or the CA3 region of the hippocampus, compared to corresponding WTs. Further analysis revealed a difference in the way the *Sorcs2*$^{-/-}$ mice responded to the novel environment when compared to the WT mice, with the *Sorcs2*$^{-/-}$ mice failing to acquire addition DSBs on exposure to exploratory task. No interaction was identified between genotype and the extent of recovery from the novel environment.

It is unlikely that the higher level of DSBs in the *Sorcs2*$^{-/-}$ mice reflect the hyperactivity phenotype in an open field test previously described for these mice[6], as their behaviour and activity levels in the home cage did not differ from those of corresponding WT mice (Dr Ditte Olsen, personal correspondence). Meanwhile, the inability of the *Sorcs2*$^{-/-}$ mice to acquire additional breaks during the exploratory activity could be perhaps linked to the hippocampal LTP and LTD deficits previously reported for these mice[5]. However, it is surprising that there was no significant difference in the extent of DNA damage in the CA region of the hippocampus between the genotypes in the home cage. *Sorcs2* is highly expressed in the CA2 region[24], and recently published findings have implicated the receptor in protecting this region from epilepsy-induced neuronal loss[22]. In mice, SorCS2 was found to be strongly upregulated in the CA2 region following induction of status epilepticus, and loss of SorCS2 was associated with increased oxidative stress and neuronal death[22].

Despite the fact that results in the WT mice that matched the *Sorcs2*$^{-/-}$ mice replicated the findings previously reported by Suberbielle *et al.* (2013)[107], the second batch of WT mice that matched the *Sorl1*$^{-/-}$ mice did not show the same pattern of transient γH2A.X foci formation following exploration of a novel environment. Moreover, these mice displayed higher levels of DSBs when compared to the WT mice from the *Sorcs2*$^{-/-}$ batch. Perhaps the most likely explanation for this surprising data is changes in the animal facility prior to the experiment, which were subsequently noted by our collaborators Dr Ditte Olsen and Mathias Kaas Ollendorff (personal communication).

There are some caveats to the experiment performed on the *Sorcs2*$^{-/-}$ mice and matched WT controls, including the small sample size (n=3) and some differences in the way the novel environment protocol was conducted compared to Suberbielle *et al.* (2013)[107]. In the study described here, WT and *Sorcs2*$^{-/-}$ mice were exposed to the novel environment individually as opposed to together with their cage-mates. Thus, it might be speculated that the increased DSB formation observed in the WT mice

belonging to the 'novel environment' group occured as a result of a stress response to being exposed to novelty on their own, rather than the novel environment itself. However, these seems unlikely to be the case as adrenalectomised mice showed similar increase in the number of DSB-positive neurons following exploration of a novel environment compared to sham-operated controls[107]. Moreover, glucocorticoids have been recently shown to upregulated *SORCS2* expression *in vitro*[23]. Thus, if the breaks detected in the 'novel environment' group were the result of corticosterone or the release of other stress response factors, it is highly likely that knocking out *Sorcs2* would have had an impact on this process.

Despite the caveats described above, the increased levels of DNA DSB formation and the inability of *Sorcs2*$^{-/-}$ mice to acquire new breaks upon exploratory activity constitute interesting findings that would need to be further validated in a bigger sample set. Due to time limitations, I was not able to do this during my PhD. However, I was able to investigate the effect of knocking out *SORCS2* on DNA DSB formation *in vitro*, as well as to explore possible mechanisms underlying this phenotype.

Aim 2: Assessing the effect of knocking out *SORCS2* on DNA DSB formation and potential underlying mechanisms *in vitro*

The second aim of my project comprised of three parts:

1) Knocking out *SORCS2* in the human neuronal cell line LUHMES, using CRISPR/Cas9 genome editing;
2) Assessing the effect of knocking out *SORCS2* on DNA DSB formation;
3) Investigating possible mechanisms underlying this effect.

Part 1: Knocking out *SORCS2* in LUHMES cells

Given their quick differentiation into a homogenous population of post-mitotic neurons, with spontaneous electric activity recorded after 10-12 days of differentiation[133], LUHMES constituted a valid system for studying the role of SORCS2 in DSB formation. Using qRT-PCR, I was able to show that all members of the sortilin receptor family, except for *SORCS3*, are expressed in the LUHMES cell line in what

appears to be a developmentally regulated manner, with their expression peaking up around day 10 and remaining stable from there on. Using CRISPR/Cas9 genome editing, I was able to knockout *SORCS2* by independently targeting exon 1 and exon 3 of its canonical, full-length isoform.

Part 2: Assessing the effect of knocking out *SORCS2* on DNA DSB formation

There was no significant difference in the number of DSBs (identified as the foci where γH2A.X and 53BP1 signal co-localised) between untreated WT (n=9) and *SORCS2* KO (n=9) neurons (day 14). Conversely, when treated with etoposide, neurons (day 14) derived from the same *SORCS2* KO clones (n=9) displayed a significantly higher levels of γH2A.X/53BP1-positive foci per nucleus compared to the WT controls (n=9). There was no difference in the number of DNA DSBs per nucleus between the *SORCS2* KO clones generated by targeting exon 1 (n=4) or exon 3 (n=5), both with and without etoposide treatment. This suggested that the increased DSB formation, observed in the *SORCS2* KO lines following treatment with etoposide, is a consequence of the mutation introduced rather than a result of CRISPR off-target activity.

The lack of significant difference in the number of DSBs per nucleus between untreated WT and *SORCS2* KO clones, as opposed to the elevated levels of DSB formation detected in the latter, following treatment with etoposide, may be explained with the low and highly variable levels of DNA damage observed in the untreated cells. In post-mitotic neurons, etoposide acts by preventing the re-ligation of DNA DSBs induced by TopoIIβ activity[162]. Therefore, treatment with etoposide would lead to the accumulation of TopoIIβ-dependent DSBs, providing a more reliable measure of all events that took place over the course of the treatment. Bearing this in mind, the fact that I was able to detect significant difference in the number of DSBs after treatment with etoposide, but not in the untreated cells, might suggest that knocking out *SORCS2 in vitro* exacerbates the rate of TopoIIβ-dependent DNA DSB formation, but has no effect on their repair. Moreover, this observation supports the findings in the *Sorcs2*$^{-/-}$ mice, which showed increased DSB levels, but no recovery deficits following the exploratory task.

Knocking out *SORCS2* was also associated with reduced neuronal viability at day 14, but not at day 6 of differentiation. This result is in line with the recently reported neuroprotective role of SORCS2[22], as well as with my early findings that in LUHMES neurons *SORCS2* expression peaks after day 10 of differentiation. However, bearing in mind the results from the untreated cells that showed no difference in the number of DSBs per nucleus between the two genotypes, it is unlikely that the reduced viability detected in the *SORCS2* KO clones is the result of accumulated DNA damage.

Both the increase in DSB formation following etoposide treatment (n=9), as well as the reduced viability observed in *SORCS2* KO day 14 neurons (n=7) constitute preliminary results from single experiments, which, due to technical problems and time constraints, I was unable to repeat. Thus, further experiments are required for these findings to be confirmed.

Part 3: Investigating possible mechanisms underlying the effect of knocking out *SORCS2* on DNA DSB formation

Stimulating WT LUHMES neurons (day 14) with NMDA led to a significant increase in the number of DSBs per nucleus, in line with previous work reporting increased γH2A.X foci formation upon NMDA application in mouse primary neurons[115,118]. I did not, however, observe any difference in the extent of γH2A.X/53BP1-positive foci formation in NMDA-treated *SORCS2* KO clones compared to WT controls. This experiment also suffers from the caveats of small sample size (n=3). Unfortunately, due to time limitations and technical issues with cell culture plasticware, I was not able to repeat it on a larger number of clones.

Given that NMDA-induced DSBs occur within the same regions of the genome as the ones triggered by etoposide[118], if the lack of difference between NMDA-treated *SORCS2* KO and WT clones was confirmed in subsequent experiments, this would suggest that the observed increase in etoposide-induced DSBs is unlikely to be due to dysregulated NMDAR signalling. Such a result would be surprising given the role of SorCS2 in the trafficking of the NMDAR subunit NR2A[51], shown to be required for DSB repair[107]. A possible explanation for such a discrepancy would be that the role of SorCS2 in the trafficking of the NR2A subunit was demonstrated in medium spiny neurons (MSNs) of the mouse striatum[51]. MSNs constitute a specific type of

GABAergic inhibitory neurons, while LUHMES are a homogenous population of human dopaminergic neurons.

Unfortunately, I was unable to measure $A\beta_{42}$ reliably in LUHMES as the levels detected were below the assay lower limit. Increased Aβ levels have been linked to increased DSB formation both *in vivo* and *in vitro*[107], and knockdown of *SORCS2* in HEK293 cells has been associated with increased APP amyloidogenic processing[46]. Therefore, elevated Aβ levels could constitute a possible explanation of the increased number of etoposide-induced DNA breaks observed in the knockout lines. Thus, future analysis could include using a more sensitive ELISA assay and/or measuring $A\beta_{40}$ and sAPPβ levels, as both metabolites of amyloidogenic APP processing have been previously detected in LUHMES neurons[134].

Surprisingly, there was no difference in the number of etoposide-induced DNA DSBs in *Sorcs2*$^{-/-}$ mouse primary hippocampal neurons (DIV12) compared to control WT cultures. In fact, looking at the absolute number of DNA breaks per nucleus, I observed a decrease in the number of DSBs both in untreated and etoposide-treated *Sorcs2*$^{-/-}$ neurons. Due to problems with mice breeding, neuronal culture and the short nature of my visit to Aarhus University, Denmark, I was able to perform this experiment only on three independent mouse cultures. It is thus possible that the small sample size led to insufficient power to detect a difference. Therefore, further experiments would be required to assess if the observed decrease represents a true result.

If the findings of decreased DSBs in the *Sorcs2*$^{-/-}$ mouse primary neurons were confirmed with subsequent experiments, there are a few possible explanations for the discrepancy with the findings of increased DSB formation detected in the dentate gyrus of SorCS2-deficient mice. One explanation could be that the primary cultures are less mature than the hippocampal neurons of an adult mouse and the observed discrepancy is a reflection of differences in experience of the adult animal versus neurons taken from a P0 pup. Alternatively, this discrepancy might suggest that the elevated levels of DSBs in the *Sorcs2*$^{-/-}$ mice are the result of altered input to the hippocampus, rather than being specific to hippocampal neurons. This possibility is supported by the increased number of DNA breaks detected in the etoposide-treated *SORCS2* KO LUHMES neurons. Therefore, it would be interesting to examine how SorCS2 deficiency affects the extent of DSB formation in brain areas receiving dopaminergic input. One such area could be the frontal cortex, which receives

dopaminergic input from the VTA and has been shown to be hyperinnervated in *Sorcs2*$^{-/-}$ mice[6]. Having said that, I did not detect a significant difference in the extracellular levels of dopamine, secreted by the *SORCS2* KO LUHMES cells. Thus, despite the high levels of variation observed in the KO clones, it is perhaps unlikely that altered dopamine levels underlie the increase in DSB formation. However, future experiments assessing aspects of dopamine metabolism and turnover might constitute an area of interest.

Aim 3: Introducing the EOAD-associated rare *SORL1* missense mutation G508S and examining its effect on *SORL1* expression and Aβ production

In addition to examining the effect of knocking out *SORCS2* on DSB formation, in the beginning of my PhD I was also planning to investigate this phenotype in SorLA-deficient mice and *SORL1* KO LUHMES neurons. Unfortunately, as discussed above and more extensively in chapter 3, I was unable to assess DSB formation in *Sorl1*$^{-/-}$ mice before and after exploratory activity, as well as following recovery from it. Subsequently, during my visit to Aarhus University, Denmark, I also aimed at investigating the levels of DNA DSB formation in primary hippocampal neurons derived from *Sorl1*$^{-/-}$ pups (P0). However, due to problems with breeding the *Sorl1*$^{-/-}$ mice, I was unable to perform these experiments.

Knocking out *SORL1* in LUHMES cells using CRISPR/Cas9 genome editing also proved to be difficult to achieve. *SORL1* encodes for a large number of transcripts, the majority of which start after exon 25 of the full-length *SORL1* transcript. Previous work in the lab performed by Susan Anderson has shown that in addition to the canonical full-length isoform, LUHMES express several shorter isoforms whose 5' ends lie within the second half of the 48 exons. Based on this, I decided to knockout the gene by targeting the first exon common for all *SORL1* transcripts- exon 31. No *SORL1* KO clones were identified following two attempts of knocking out the gene using a gRNA targeting this region. With hindsight, these results were not surprising given that exon 31 is located far away from the beginning of the full-length transcript. Thus, frameshift mutations introduced at this region were more likely to generate truncated versions of the long *SORL1* isoform, rather than activate NMD pathways. A more efficient approach of knocking out *SORL1* would have been generating a whole

gene deletion by introducing two gRNAs, targeting either ends of the gene. This would result in the formation of two DSBs at the targeted locus, which would be repaired via NHEJ pathway, deleting the intervening DNA segment.

In parallel to my attempts to knock out the *SORL1* gene, I aimed to introduce one of the many rare EOAD-associated *SORL1* variants- G508S, in LUHMES cells using CRISPR/Cas9 genome editing. This specific variant was chosen for many reasons (discussed in chapter 6), amongst which was its location. G508S constitutes a missense mutation occurring at the last nucleotide of exon 10 and thus affecting a 5' splice region. Therefore, it was highly likely that it might affect *SORL1* splicing. Additionally, any indels in this region could have potentially altered splicing and knocked out the full-length *SORL1* transcript.

I was unable to identify any *SORL1* KO clones amongst the clones that were homozygous for a frameshift mutation within the targeted region. However, two of these clones showed reduced SORLA protein levels and were identified as knockdowns. I also generated four G508S heterozygous mutants and three compound heterozygotes, carrying the G508S mutation on one allele and an indel on the other. However, only one G508S heterozygous mutant was revived successfully.

RT-PCR performed on cDNA derived from the G508S heterozygous clone showed no changes in *SORL1* splicing. However, I detected lower *SORL1* mRNA levels in both the G508S heterozygous and the three compound heterozygous clones, which did not translate into an alteration at the protein level.

As mentioned above, these experiments were performed on only one clone. Thus, if I was to continue this project, first I would create more mutant lines to ensure that any phenotypes identified are the result of the mutation itself. Provided more lines carrying the G508S mutation were generated, I would continue by repeating the experiments described above, but given that I had shown that the mutation does not impact splicing, I would assess whether it affects APP processing and Aβ production. This could be achieved by either measuring the amounts of sAPPα, sAPPβ and Aβ produced by the mutant clones using ELISA, or, alternatively, by measuring levels of secretase activity as described by Reitz *et al.* (2013)[46]. Unfortunately, I was unable to measure reliably the levels of extracellular $A\beta_{42}$ produced by both WT and *SORL1* mutant clones. This suggests that a more sensitive system than the one used in this study would be required if this experiment is to be repeated in the future. Moreover, since the G508S variant is in a close proximity to the SORLA binding site for Aβ,

assessing how the mutation affects the receptor's ability to direct Aβ molecules for lysosomal degradation might also be of interest. Finally, if the G508S mutation was found to promote Aβ accumulation, investigating its effect on DSB formation might also be of interest.

Regardless of all caveats of this part of my project, I was successful in introducing the rare EOAD-associated *SORL1* missense mutation- G508S. I also showed, based on the one cloned examined, that there is no evidence that the G508S mutation affects *SORL1* splicing. Although I was unable to further analyse the effect of this mutation on SORLA function, the optimisation of the CRISPR process for knocking in mutations at this position can be used to create more independently-derived G508S mutant lines or adapted to introduce other EOAD-associated variants in the future.

Other Future Directions

As discussed in chapter 1, SORCS2 has been recently shown to play a role in the neuronal defence against oxidative stress[22]. It does so by facilitating the trafficking of EAAT3 to the cell surface and thus the import of cysteine, required for the synthesis of the ROS scavenger glutathione. Loss of SorCS2 in mice has been associated with elevated levels of the oxidative stress DNA marker 8-OHdG and cell death in the dentate gyrus and the CA2/3 region of the hippocampus following PTZ kindling[22]. Despite the disagreement in the literature regarding the link between oxidative stress and γH2A.X foci formation, it would be interesting to investigate whether the *Sorcs2*$^{-/-}$ mice in our study, as well as the *SORCS2* KO LUHMES clones, exhibit higher 8-OHdG levels. If so, future work may also involve co-staining for γH2A.X and 8-OHdG and establishing the levels of overlap between the two markers of DNA damage in WT and SorCS2-deficient mice and/or cells. Examining the response of *SORCS2* KO LUHMES neurons to oxidative stress could also constitute an area of interest. Stressing the cells with hydrogen peroxide and then measuring 8-OHdG levels and/or expression of genes involved in oxidative stress defence pathways might be one way of approaching this. It would also be interesting to test whether inducing oxidative stress would exacerbate the difference in neuronal viability I detected in day 14 *SORCS2* KO LUHMES neurons. In addition, given the recently identified role of SORCS2 in the cellular response to alcohol[23], following experiments looking at the consequences of treating *SORCS2* KO neurons with alcohol might prove to be illuminating, as well.

Madabushi *et al.* (2015)[118] showed that TopoIIβ-dependent DSBs promote the expression of early response genes, including *Fos*, *FosB*, *Npas* and *Egr1*. If knocking out *SORCS2* exacerbates DNA DSB formation *in vivo* and *in vitro*, this poses the question how such an increase in DSBs affects the expression levels of these genes and what the downstream consequences are, in terms of neuronal function and survival. In rat ganglion neurons, accumulation of unrepaired DSBs has been associated with transcriptional silencing, potentially as a mechanism of preventing aberrant mRNA synthesis and protein production[187]. Thus, it would be interesting to investigate whether the increased DSB formation in the *SORCS2* KO clones leads to a reduced expression of early-response genes.

In conclusion, both increased DNA damage and members of the sortilin gene family have been implicated in the aetiology of many brain disorders. The work completed in this project constitutes the first evidence for a potential role of the sortilins in the generation of correct, physiological DSBs and protection against aberrant DSB formation. Future work investigating this link further might provide us with valuable knowledge on the cellular mechanisms underlying neurodegenerative and psychiatric conditions.

Appendices

Appendix A- Summary of the actual count number per group/ experiment and the results of the statistical analyses described in Chapter 3

Table A1 Counts obtained per mouse per experimental group from the preliminary experiment described in 3.4.1. Mean percentage of γH2A.X-positive neurons in the dentate gyrus and the CA3 region of the hippocampus of WT and Sorcs2$^{-/-}$ mice belonging to one of the three experimental groups- home cage, Novel E and recovery.

Condition	**Dentate gyrus**		**CA3 region**	
	WT	***Sorcs2*$^{-/-}$**	**WT**	***Sorcs2*$^{-/-}$**
Home cage	24.439	32.374	33.444	36.807
	22.870	26.864	30.358	35.966
	22.646	29.107	31.524	35.087
Novel E	37.276	35.535	36.763	42.089
	33.242	28.459	38.215	36.374
	38.680	39.636	46.870	39.252
Recovery	25.395	27.890	27.927	40.070
	25.156	26.504	33.674	44.237
	29.118	26.381	33.630	33.450

Table A2 Assessing the effect of the environment on the formation of DSBs in the beginning of the hippocampus of WT and Sorcs2$^{-/-}$ mice. Summary of the results obtained from post-hoc tests as part of the parametric One-way ANOVA analysis performed in 3.4.1. 'E' stands for environment, 'HC'- home cage, 'NE'- novel environment and 'R'- recovery.

Brain region	Environment		Mean percentage		Mean Difference	t-value	p-value
	E1	E2	E1	E2			
Wild type mice							
Dentate gyrus	HC	NE	23.32	36.4	-13.08	10.54	0.0007
	HC	R	23.32	26.56	-3.237	2.609	0.2344
	NE	R	36.4	26.56	9.843	7.932	0.0033
CA3	HC	NE	31.78	40.62	-8.84	4.035	0.0653
	HC	R	31.78	31.74	0.03232	0.01475	>0.9999
	NE	R	40.62	31.74	8.873	4.049	0.0645
***Sorcs2*$^{-/-}$ mice**							
Dentate gyrus	HC	NE	29.45	34.54	-5.095	2.406	0.2796
	HC	R	29.45	26.92	2.523	1.192	0.6925
	NE	R	34.54	26.92	7.618	3.598	0.0965
CA3	HC	NE	35.95	39.24	-3.285	1.588	0.5358
	HC	R	35.95	39.25	-3.299	1.595	0.5333
	NE	R	39.24	39.25	-0.01404	0.00679	>0.9999

Table A3 Counts obtained per experimental group per mouse from the preliminary experiment described in 3.4.2. Mean percentage of γH2A.X-positive neurons in the dentate gyrus and the CA3 region of the hippocampus of WT and Sorl1$^{-/-}$ mice belonging to one of the three experimental groups- home cage, Novel E and recovery.

Condition	Dentate gyrus		CA3 region	
	WT	***Sorl1*$^{-/-}$**	WT	***Sorl1*$^{-/-}$**
Home cage	29.073	48.225	67.192	63.553
	47.4908	32.427	65.989	71.743
	51.928	15.058	59.661	53.879
Novel E	51.260	27.098	48.109	11.030
	35.891	32.332	41.942	65.862
	19.551	11.513	65.352	61.598
	4.231		70.866	64.409
Recovery	55.201	29.702	56.192	52.608
	22.174	62.687	69.061	49.016
	45.022	32.625	52.257	68.358
		54.598		65.343

Table A4 Assessing the effect of the environment on the formation of DSBs in the beginning of the hippocampus of WT and Sorl1$^{-/-}$ mice. Summary of the results obtained from post-hoc tests as part of the parametric One-way ANOVA analysis performed in 3.4.2. 'E' stands for environment, 'HC'- home cage, 'NE'- novel environment and 'R'- recovery.

Brain region	Environment		Mean percentage		Mean Difference	t-value	p-value
	E1	E2	E1	E2			
Wild type mice							
Dentate gyrus	HC	NE	42.83	27.73	15.1	1.612	0.5221
	HC	R	42.83	40.8	2.032	0.2029	0.9887
	NE	R	27.73	40.8	-13.07	1.395	0.6077
CA3	HC	NE	64.28	56.57	7.713	1.376	0.6155
	HC	R	64.28	59.17	5.11	0.8527	0.8231
	NE	R	56.57	59.17	-2.603	0.4643	0.9428
***Sorl1*$^{-/-}$ mice**							
Dentate gyrus	HC	NE	31.9	23.65	8.255	0.9526	0.7856
	HC	R	31.9	44.9	-13	1.604	0.5253
	NE	R	23.65	44.9	-21.26	2.622	0.2215
CA3	HC	NE	63.06	50.72	12.33	1.282	0.6515
	HC	R	63.06	58.83	4.227	0.4394	0.9485
	NE	R	50.72	58.83	-8.107	0.9102	0.8010

Table A5 Counts obtained per experimental group per mouse from the experiment described in 3.5. Mean percentage of γH2A.X-positive neurons in the dentate gyrus, the CA2 and CA3 region of the hippocampus of WT and Sorcs2$^{-/-}$ mice belonging to one of the three experimental groups- home cage, Novel E and recovery.

Condition	Dentate gyrus		CA2 region		CA3 region	
	WT	*Sorcs2*$^{-/-}$	WT	*Sorcs2*$^{-/-}$	WT	*Sorcs2*$^{-/-}$
Home cage	31.428	46.045	25.897	57.113	28.139	57.732
	31.503	42.373	38.609	34.846	40.152	45.332
	39.848	50.542	38.472	42.416	53.748	53.585
Novel E	59.427	57.852	57.033	54.084	54.013	53.965
	59.682	45.817	53.763	49.346	56.323	51.092
	55.031	52.367	62.204	55.350	63.779	50.958
Recovery	43.440	45.977	23.859	35.745	29.263	39.485
	42.683	45.603	49.860	47.797	49.092	56.438
	39.069	43.874	39.301	40.810	37.072	42.087

Table A6 Assessing the effect of the environment on the formation of DSBs in the hippocampus of WT and Sorcs2$^{-/-}$ mice. Summary of the results obtained from the non-parametric ANOVA-type analysis performed in 3.4.3. 'E' stands for environment, 'HC'- home cage, 'NE'- novel environment and 'R'- recovery.

Brain region	**Environment**		**Rank means**		**Rank Mean Difference**	**t-value**	**p-value**
	E1	**E2**	**E1**	**E2**			
Wild type mice							
Dentate gyrus	HC	NE	5.44	13.56	-8.12	34.38	4.53×10^{-9}
	HC	R	7.78	11.22	-3.44	3.65	0.056
	NE	R	12.33	6.67	5.66	22.82	1.78×10^{-6}
CA2	HC	NE	3.50	9.50	-6	18.78	1.46×10^{-5}
	HC	R	6.17	6.83	-0.66	0.08	0.774
	NE	R	9.00	4.00	5	6.52	0.011
CA3	HC	NE	4.17	8.83	-4.66	4.45	0.035
	HC	R	7.00	6.00	1	0.12	0.734
	NE	R	9.00	4.00	5	7.50	0.006
***Sorcs2*$^{-/-}$ mice**							
Dentate gyrus	HC	NE	8.11	10.89	-2.78	1.52	0.218
	HC	R	9.56	9.44	0.12	0.04	0.835
	NE	R	11.56	7.44	4.12	2.91	0.088
CA2	HC	NE	5.83	7.17	-1.34	0.32	0.574
	HC	R	7.17	5.83	1.34	0.44	0.508
	NE	R	7.83	5.17	2.66	3.20	0.074
CA3	HC	NE	6.67	6.33	0.34	0.02	0.889
	HC	R	7.67	5.33	2.34	1.08	0.299
	NE	R	8.17	4.83	3.34	2.86	0.091

Table A7 Assessing the effect of the genotype-environment interaction on the formation of DSBs. Summary of the results obtained from the non-parametric ANOVA-type analysis performed in 3.4.3. fitting in the genotype by interaction. 'E' stands for environment, 'HC'- home cage, 'NE'- novel environment and 'R'- recovery.

Brain region	**Environment**		**Rank means**				**t-value**	**p-value**
	E1	**E2**	**WTxE1**	**WTxE2**	**KOxE1**	**KOxE2**		
Dentate gyrus	HC	NE	8.22	26.44	16.89	22.44	10.50	0.001
	NE	R	25.67	12.56	21.22	14.56	1.66	0.198
CA2	HC	NE	19.00	11.67	14.17	12.33	5.35	0.021
	NE	R	7.50	14.00	9.50	14.50	1.99	0.158
CA3	HC	NE	17.83	13.00	12.50	14.17	3.87	0.049
	NE	R	6.67	14.50	10.00	15.17	2.19	0.139

Table A8 Assessing the effect of the environment on the formation of DSBs. Summary of the results obtained from the non-parametric ANOVA-type analysis performed in 3.4.3. fitting in the genotype by interaction. 'E' stands for environment, 'HC'- home cage, 'NE'- novel environment and 'R'- recovery.

Brain region	**Environment**		**Rank means**		**t-value**	**p-value**
	E1	**E2**	**E1**	**E2**		
Dentate gyrus	HC	NE	12.56	24.44	37.00	1.18×10^{-9}
	HC	R	16.50	20.50	5.97	0.015
	NE	R	23.44	13.56	15.63	7.71×10^{-5}
CA2	HC	NE	8.42	16.58	11.12	0.001
	HC	R	12.50	12.50	0.00	1.000
	NE	R	16.50	8.50	10.41	0.001
CA3	HC	NE	9.83	15.17	3.24	0.072
	HC	R	13.75	11.25	0.54	0.462
	NE	R	16.67	8.33	10.33	0.001

Table A9 Assessing the effect of the genotype on the formation of DSBs. Summary of the results obtained from the non-parametric ANOVA-type analysis performed in 3.4.3. fitting in the genotype by interaction. 'E' stands for environment, 'HC'- home cage, 'NE'- novel environment and 'R'- recovery.

Brain region	**Environment**		**Rank means**		**t-value**	**p-value**
	E1	**E2**	**Wild type**	***Sorcs2*$^{-/-}$**		
Dentate gyrus	HC	NE	17.33	19.67	1.43	0.233
	HC	R	15.39	21.61	14.45	1.44×10^{-4}
	NE	R	19.11	17.89	0.24	0.625
CA2	HC	NE	12.08	12.92	0.12	0.734
	HC	R	10.50	14.50	1.68	0.195
	NE	R	13.25	11.75	0.37	0.545
CA3	HC	NE	12.25	12.75	0.03	0.866
	HC	R	9.42	15.58	3.29	0.070
	NE	R	12.75	12.25	0.04	0.847

Appendix B- Supplementary figures

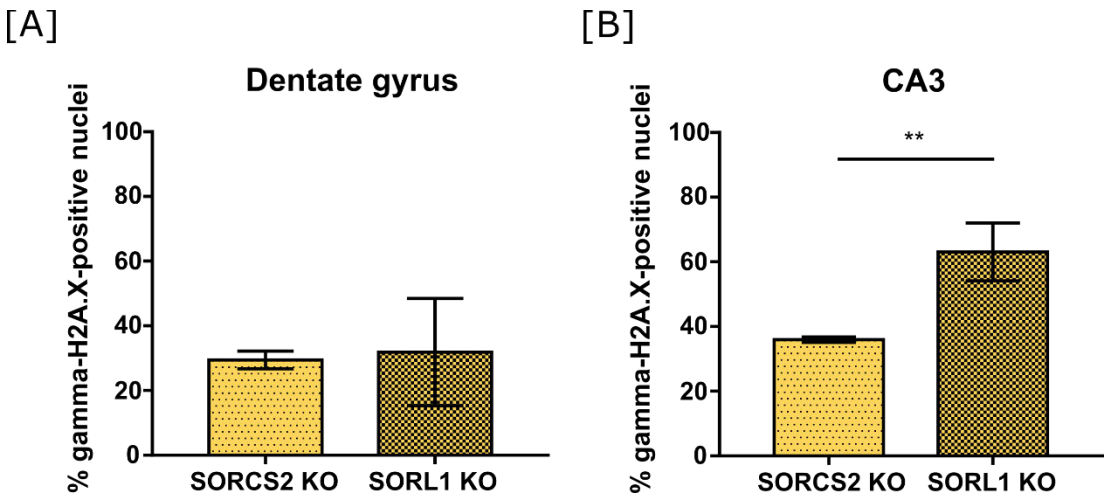

*Figure B1 Comparison of the percentage of γH2A.X-positive nuclei detected in the dentate gyrus [A] and the CA3 region of the hippocampus [B] of Sorcs2$^{-/-}$ and Sorl1$^{-/-}$ mice. Error bars represent means ± SD; unpaired t-test; '**' indicates p<0.01.*

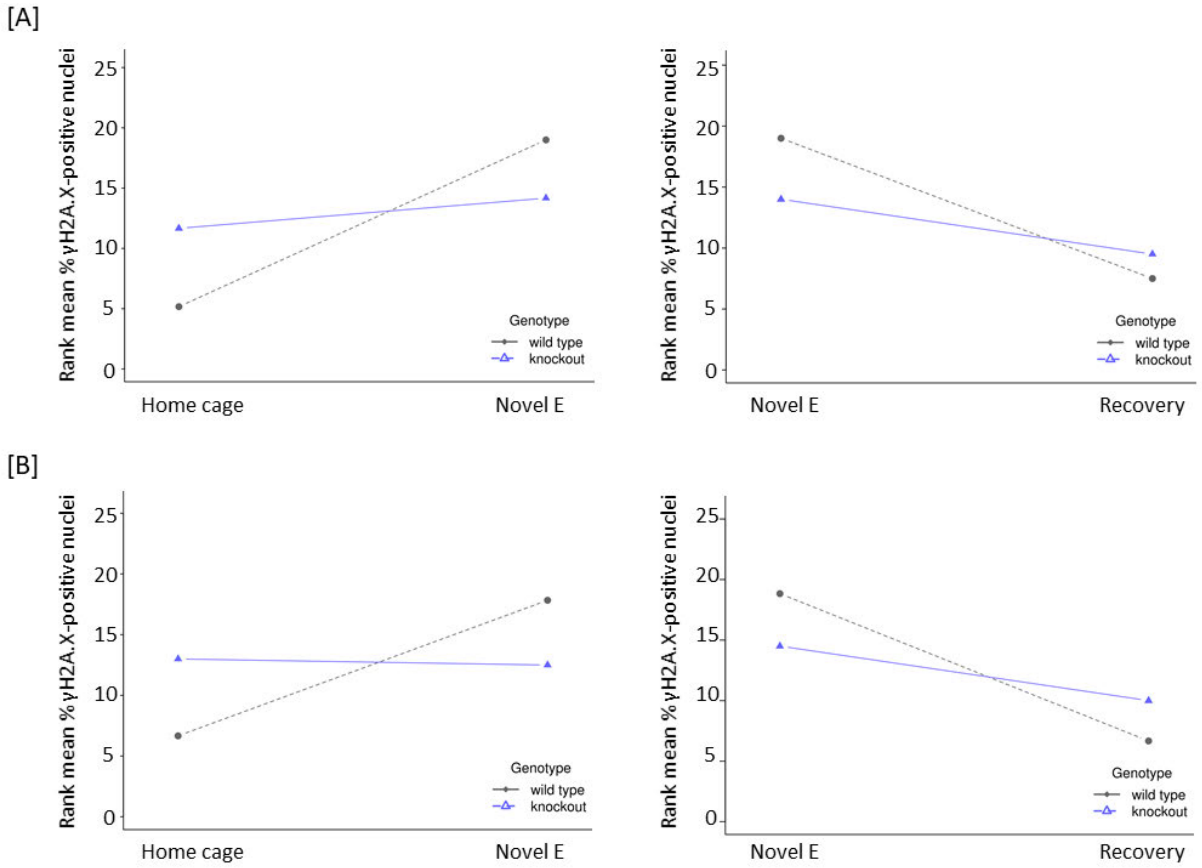

Figure B2 Comparison of the rate of DSB formation (left) and repair (right) in the CA2 [A] and CA3 [B] regions in WT and Sorcs2$^{-/-}$ mice. n=3, non-parametric ANOVA-type statistic.

Bibliography

1. Hermey, G. The Vps10p-domain receptor family. *Cell. Mol. Life Sci.* **66**, 2677–2689 (2009). doi: 10.1007/s00018-009-0043-1
2. Glerup, S., Nykjaer, A. & Vaegter, C. B. Sortilins in Neurotrophic Factor Signaling. *Handb. Exp. Pharmacol.* **220**, 165–189 (2014). doi:10.1007/978-3-642-45106-5_7
3. Glerup, S. *et al.* SorLA Controls Neurotrophic Activity by Sorting of GDNF and Its Receptors GFRα1 and RET. *CellReports* **3**, 186–199 (2013). doi: 10.1016/j.celrep.2012.12.011
4. Deinhardt, K. *et al.* Neuronal growth cone retraction relies on proneurotrophin receptor signaling through Rac. *Sci. Signal.* **4**, ra82 (2011). doi: 10.1126/scisignal.2002060.
5. Glerup, S. *et al.* SorCS2 is required for BDNF-dependent plasticity in the hippocampus. *Mol. Psychiatry* **21108**, 1740–1751 (2016). doi: 10.1038/mp.2016.108
6. Glerup, S. *et al.* SorCS2 regulates dopaminergic wiring and is processed into an apoptotic two-chain receptor in peripheral glia. *Neuron **82**, 1074-87* (2014). doi:10.1016/j.neuron.2014.04.022
7. Mehmedbasic, A. *et al.* SorLA Complement-type Repeat Domains Protect the Amyloid Precursor Protein against Processing. *J. Biol. Chem.* **290**, 3359-3376 (2014). doi:10.1074/jbc.M114.619940
8. Schmidt, V., Subkhangulova, A. & Willnow, T. E. Sorting receptor SORLA: cellular mechanisms and implications for disease. *Cell. Mol. Life Sci.* **74**, 1475–1483 (2017). doi: 10.1007/s00018-016-2410-z
9. Andersen, O. M., Rudolph, I.-M. & Willnow, T. E. Risk factor SORL1: from genetic association to functional validation in Alzheimer's disease. *Acta Neuropathol.* **132**, 653–665 (2016). doi: 10.1007/s00401-016-1615-4
10. Willnow, T. E., Petersen, C. M. & Nykjaer, A. VPS10P-domain receptors — regulators of neuronal viability and function. *Nat. Rev. Neurosci.* **9**, 899–909 (2008). doi: 10.1038/nrn2516
11. Schmidt, V. *et al.* SorLA/LR11 Regulates Processing of Amyloid Precursor Protein via Interaction with Adaptors GGA and PACS-1. *J. Biol. Chem.* **282**:32956-32964 (2007). doi:10.1074/jbc.M705073200
12. Fjorback, A. W. *et al.* Retromer binds the FANSHY sorting motif in SorLA to

regulate amyloid precursor protein sorting and processing. *J. Neurosci.* **32**, 1467–1480 (2012). doi: 10.1523/JNEUROSCI.2272-11.2012

13. Jacobsen, L. *et al.* The sorLA cytoplasmic domain interacts with GGA1 and -2 and defines minimum requirements for GGA binding. *FEBS Lett.* **511**, 155–158 (2002). doi: 10.1016/s0014-5793(01)03299-9
14. Huang, T. Y. *et al.* SNX27 and SORLA Interact to Reduce Amyloidogenic Subcellular Distribution and Processing of Amyloid Precursor Protein. *J. Neurosci.* **36**, 7996–8011 (2016). doi: 10.1523/JNEUROSCI.0206-16.2016
15. Dumanis, S. B. *et al.* Distinct Functions for Anterograde and Retrograde Sorting of SORLA in Amyloidogenic Processes in the Brain. *J. Neurosci.* **35**,12703-12713 (2015). doi:10.1523/JNEUROSCI.0427-15.2015
16. Hermey, G., Sjøgaard, S. S., Petersen, C. M., Nykjær, A. & Gliemann, J. Tumour necrosis factor α-converting enzyme mediates ectodomain shedding of Vps10p-domain receptor family members. *Biochem. J.* **395**, 285–293 (2006). doi: 10.1042/BJ20051364
17. Böhm, C. *et al.* SorLA Signaling by Regulated Intramembrane Proteolysis. *J. Biol. Chem.* **281**, 14547–14553 (2006). doi: 10.1074/jbc.M601660200
18. Hermans-Borgmeyer, I. *et al.* Unique expression pattern of a novel mosaic receptor in the developing cerebral cortex. *Mech. Dev.* **70**, 65–76 (1998). doi: 10.1016/s0925-4773(97)00177-9
19. Rezgaoui, M. *et al.* Identification of SorCS2, a novel member of the VPS10 domain containing receptor family, prominently expressed in the developing mouse brain. *Mech. Dev.* **100**, 335-338 (2001). doi:10.1016/S0925-4773(00)00523-2
20. Hermey, G. *et al.* The three sorCS genes are differentially expressed and regulated by synaptic activity. *J. Neurochem.* **88**, 1470-6 (2004). doi:10.1046/j.1471-4159.2004.02286.x
21. Oetjen, S., Mahlke, C., Hermans-Borgmeyer, I. & Hermey, G. Spatiotemporal Expression Analysis of the Growth Factor Receptor SorCS3. *J. Comp. Neurol.* **522**, 3386-402 (2014). doi:10.1002/cne.23606
22. Malik, A. R. *et al.* SorCS2 Controls Functional Expression of Amino Acid Transporter EAAT3 and Protects Neurons from Oxidative Stress and Epilepsy-Induced Pathology . *Cell Rep.* **26**, 2792-2804 (2019). doi: 10.1016/j.celrep.2019.02.027
23. Smith, A. H. *et al.* Risk Locus Identification Ties Alcohol Withdrawal Symptoms

to *SORCS2*. *Alcohol. Clin. Exp. Res.* **42**, 2337–2348 (2018). doi: 10.1111/acer.13890

24. Hermey, G. *et al.* The three sorCS genes are differentially expressed and regulated by synaptic activity. *J. Neurochem.* **88**, 1470–1476 (2004). doi: 10.1046/j.1471-4159.2004.02286.x
25. Sarret, P. *et al.* Distribution of NTS3 receptor/sortilin mRNA and protein in the rat central nervous system. *J. Comp. Neurol.* **461**, 483–505 (2003). doi: 10.1002/cne.10708
26. Motoi, Y. *et al.* Neuronal localization of a novel mosaic apolipoprotein E receptor, LR11, in rat and human brain. *Brain Res.* **833**, 209–215 (1999). doi: 10.1016/s0006-8993(99)01542-5
27. Westergaard, U. B. *et al.* SorCS3 does not require propeptide cleavage to bind nerve growth factor. *FEBS Lett.* **579**, 1172–1176 (2005). doi: 10.1016/j.febslet.2004.12.088
28. Reichardt, L. F. Neurotrophin-regulated signalling pathways. *Philos. Trans. R. Soc. B Biol. Sci.* **361**, 1545 (2006). doi: 10.1098/rstb.2006.1894
29. Nykjaer, A. *et al.* Sortilin is essential for proNGF-induced neuronal cell death. *Nature* **427**, 843–848 (2004). doi: 10.1038/nature02319
30. Jansen, P. *et al.* Roles for the pro-neurotrophin receptor sortilin in neuronal development, aging and brain injury. *Nat. Neurosci.* **10**, 1449–1457 (2007). doi: 10.1038/nn2000
31. Vaegter, C. B. *et al.* Sortilin associates with Trk receptors to enhance anterograde transport and neurotrophin signaling. *Nat. Neurosci.* **14**, 54–63 (2011). doi: 10.1038/nn.2689
32. Chen, Z.-Y. *et al.* Sortilin controls intracellular sorting of brain-derived neurotrophic factor to the regulated secretory pathway. *J. Neurosci.* **25**, 6156–66 (2005). doi: 10.1523/JNEUROSCI.1017-05.2005
33. Rohe, M., Hartl, D., Fjorback, A. N., Klose, J. & Willnow, T. E. SORLA-Mediated Trafficking of TrkB Enhances the Response of Neurons to BDNF. *PLoS One* **8**, e72164 (2013). doi: 10.1371/journal.pone.0072164
34. Rohe, M. *et al.* Brain-Derived Neurotrophic Factor Reduces Amyloidogenic Processing through Control of SORLA Gene Expression. *J. Neurosci.* **29**, 15472–15478 (2009). doi: 10.1523/JNEUROSCI.3960-09.2009
35. Perl, D. P. Neuropathology of Alzheimer's disease. *Mt. Sinai J. Med.* **77**, 32–42 (2010). doi: 10.1002/msj.20157

36. Müller, U. C., Deller, T. & Korte, M. Not just amyloid: physiological functions of the amyloid precursor protein family. *Nat. Rev. Neurosci.* **18**, 281–298 (2017). doi: 10.1038/nrn.2017.29
37. Andersen, O. M. *et al.* Neuronal sorting protein-related receptor sorLA/LR11 regulates processing of the amyloid precursor protein. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 13461–13466 (2005). doi: 10.1073/pnas.0503689102
38. Spoelgen, R. *et al.* Interaction of the Cytosolic Domains of sorLA/LR11 with the Amyloid Precursor Protein (APP) and-Secretase-Site APP-Cleaving Enzyme. *J. Neurosci.* **26**, 418-428 (2006). doi:10.1523/JNEUROSCI.3882-05.2006
39. Offe, K. *et al.* The Lipoprotein Receptor LR11 Regulates Amyloid Production and Amyloid Precursor Protein Traffic in Endosomal Compartments. *J. Neurosci.* **26**, 1596-1603 (2006). doi:10.1523/JNEUROSCI.4946-05.2006
40. Burgert, T. *et al.* SORLA-dependent and -independent functions for PACS1 in control of amyloidogenic processes. *Mol. Cell. Biol.* **33**, 4308–4320 (2013). doi: 10.1128/MCB.00628-13
41. Caglayan, S. *et al.* Lysosomal Sorting of Amyloid-b by the SORLA Receptor Is Impaired by a Familial Alzheimer's Disease Mutation. *Sci. Transl. Med.* **6**, 223ra20. (2014). doi: 10.1126/scitranslmed.3007747
42. Dodson, S. E. *et al.* Loss of LR11/SORLA Enhances Early Pathology in a Mouse Model of Amyloidosis: Evidence for a Proximal Role in Alzheimer's Disease. *J. Neurosci.* **28**, 12877-12886 (2008). doi:10.1523/JNEUROSCI.4582-08.2008
43. Capsoni, S. *et al.* SorLA Deficiency Dissects Amyloid Pathology from Tau and Cholinergic Neurodegeneration in a Mouse Model of Alzheimer's Disease. *J. Alzheimer's Dis.* **33**, 357–371 (2013). doi: 10.3233/JAD-2012-121399
44. Kim, J., Basak, J. M. & Holtzman, D. M. The role of apolipoprotein E in Alzheimer's disease. *Neuron* **63**, 287–303 (2009). doi: 10.1016/j.neuron.2009.06.026
45. Yajima, R. *et al.* ApoE-isoform-dependent cellular uptake of amyloid-β is mediated by lipoprotein receptor LR11/SorLA. *Biochem. Biophys. Res. Commun.* **456**, 482–488 (2015). doi: 10.1016/j.bbrc.2014.11.111
46. Reitz, C. *et al.* Independent and epistatic effects of variants in VPS10-d receptors on Alzheimer disease risk and processing of the amyloid precursor protein (APP). *Transl. Psychiatry* **3**, e256–e256 (2013). doi: 10.1038/tp.2013.13

47. Reitz, C. *et al.* SORCS1 alters amyloid precursor protein processing and variants may increase Alzheimer's disease risk. *Ann. Neurol.* **69**, 47–64 (2011). doi: 10.1002/ana.22308
48. Carlo, A.-S. *et al.* The Pro-Neurotrophin Receptor Sortilin Is a Major Neuronal Apolipoprotein E Receptor for Catabolism of Amyloid-Peptide in the Brain. *J. Neurosci.* **33**, 358-370 (2013). doi:10.1523/JNEUROSCI.2425-12.2013
49. Gustafsen, C. *et al.* Sortilin and SorLA Display Distinct Roles in Processing and Trafficking of Amyloid Precursor Protein. *J. Neurosci.* **33**, 64-71 (2013). doi:10.1523/JNEUROSCI.2371-12.2013
50. Ferrer, I., Goutan, E., Marín, C., Rey, M. J. & Ribalta, T. Brain-derived neurotrophic factor in Huntington disease. *Brain Res.* **866**, 257–261 (2000).
51. Ma, Q. *et al.* SorCS2-mediated NR2A trafficking regulates motor deficits in Huntington's disease. *JCI Insight.* **2**, pii: 88995 (2017). doi: 10.1172/jci.insight.88995
52. Dhungel, N. *et al.* Parkinson's Disease Genes VPS35 and EIF4G1 Interact Genetically and Converge on α-Synuclein. *Neuron* **85**, 76–87 (2015). doi: 10.1016/j.neuron.2014.11.027
53. Hu, F. *et al.* Sortilin-mediated endocytosis determines levels of the frontotemporal dementia protein, progranulin. *Neuron* **68**, 654–667 (2010). doi: 10.1016/j.neuron.2010.09.034
54. Olsen, D. *et al.* Reduced Alcohol Seeking and Withdrawal Symptoms in Mice Lacking the BDNF Receptor SorCS2. *Front. Pharmacol.* **10**, 499 (2019). doi: 10.3389/fphar.2019.00499
55. Capsoni, S. *et al.* Dissecting the role of sortilin receptor signaling in neurodegeneration induced by NGF deprivation. *Biochem. Biophys. Res. Commun.* **431**, 579-585 (2013). doi: 10.1016/j.bbrc.2013.01.007
56. Moreno, S. *et al.* c. *Front. Pharmacol.* **9**, 863 (2018). doi: 10.3389/fphar.2018.00863
57. Christiansen, G. B. *et al.* The sorting receptor SorCS3 is a stronger regulator of glutamate receptor functions compared to GABAergic mechanisms in the hippocampus. *Hippocampus* **27**, 235–248 (2017). doi: 10.1002/hipo.22689
58. Breiderhoff, T. *et al.* Sortilin-Related Receptor SORCS3 Is a Postsynaptic Modulator of Synaptic Depression and Fear Extinction. *PLoS One* **8**, e75006 (2013). doi: 10.1371/journal.pone.0075006
59. Lee, J. J. *et al.* Gene discovery and polygenic prediction from a genome-wide

association study of educational attainment in 1.1 million individuals. *Nat. Genet.* **50**, 1112–1121 (2018). doi: 10.1038/s41588-018-0147-3

60. Guerra-Carrillo, B., Katovich, K. & Bunge, S. A. Does higher education hone cognitive functioning and learning efficacy? Findings from a large and diverse sample. *PLoS One* **12**, e0182276 (2017). doi: 10.1371/journal.pone.0182276
61. Cowan, R., Hurry, J. & Midouhas, E. The relationship between learning mathematics and general cognitive ability in primary school. *Br. J. Dev. Psychol.* **36**, 277–284 (2018). doi.org/10.1111/bjdp.12200
62. Davies, G. *et al.* Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. *Nat. Commun.* **9**, 2098 (2018). doi: 10.1038/s41467-018-04362-x
63. Rogaeva, E. *et al.* The neuronal sortilin-related receptor SORL1 is genetically associated with Alzheimer disease. *Nat. Genet.* **39**, 168–177 (2007). doi: 10.1038/ng1943
64. Young, J. E. *et al.* Elucidating molecular phenotypes caused by the SORL1 Alzheimer's disease genetic risk factor using human induced pluripotent stem cells. *Cell Stem Cell* **16**, 373–85 (2015). doi: 10.1016/j.stem.2015.02.004
65. Cong, L. *et al.* Association between SORL1 polymorphisms and the risk of Alzheimer's disease. *J. Integr. Neurosci.* **17**, 185–192 (2018). doi: 10.31083/JIN-170051
66. Reitz, C. *et al.* Meta-analysis of the Association Between Variants in SORL1 and Alzheimer Disease. *Arch. Neurol.* **68**, 99-106 (2011).. doi:10.1001/archneurol.2010.346
67. Lee, J. H. *et al.* The Association Between Genetic Variants in SORL1 and Alzheimer's Disease in an Urban, Multiethnic, Community-Based Cohort. *Arch Neurol* **64**, 501–506 (2007). doi: 10.1001/archneur.64.4.501
68. Kunkle, B. W. *et al.* Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. *Nat. Genet.* **51**, 414–430 (2019). doi: 10.1038/s41588-019-0358-2
69. Lambert, J.-C. *et al.* Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat. Genet.* **45**, 1452–1458 (2013). doi: 10.1038/ng.2802
70. Jansen, I. E. *et al.* Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat. Genet.* **51**, 404–413 (2019). doi: 10.1038/s41588-018-0311-9

71. Marioni, R. E. *et al.* GWAS on family history of Alzheimer's disease. *Transl. Psychiatry* **8**, 99 (2018). doi: 10.1038/s41398-018-0150-6
72. Piscopo, P. *et al.* SORL1 Gene is Associated with the Conversion from Mild Cognitive Impairment to Alzheimer's Disease. *J. Alzheimer's Dis.* **46**, 771–776 (2015). doi: 10.3233/JAD-141551
73. Pottier, C. *et al.* High frequency of potentially pathogenic SORL1 mutations in autosomal dominant early-onset Alzheimer disease. *Mol. Psychiatry* **17**, 875–879 (2012). doi: 10.1038/mp.2012
74. Campion, D., Charbonnier, C. & Nicolas, G. SORL1 genetic variants and Alzheimer disease risk: a literature review and meta-analysis of sequencing data. *Acta Neuropathol.* **138**, 173–186 (2019). doi: 10.1007/s00401-019-01991-4
75. Vardarajan, B. N. *et al.* Coding mutations in SORL1 and Alzheimer disease. *Ann. Neurol.* **77**, 215-227 (2015). doi:10.1002/ana.24305
76. Cuccaro, M. L. *et al.* SORL1 mutations in early-and late-onset Alzheimer disease. *Neurol. Genet.* **2**, e116 (2016). doi:10.1212/NXG.0000000000000116
77. Verheijen, J. *et al.* A comprehensive study of the genetic impact of rare variants in SORL1 in European early-onset Alzheimer's disease. *Acta Neuropathol.* **132**, 1–12 (2016). doi: 10.1007/s00401-016-1566-9
78. Bellenguez, C. *et al.* Contribution to Alzheimer's disease risk of rare variants in TREM2, SORL1, and ABCA7 in 1779 cases and 1273 controls. *Neurobiol. Aging* **59**, 220.e1-220.e9 (2017). doi: 10.1016/j.neurobiolaging.2017.07.001
79. Thonberg, H. *et al.* Identification and description of three families with familial Alzheimer disease that segregate variants in the SORL1 gene. *Acta Neuropathol. Commun.* **5**, 43 (2017). doi: 10.1186/s40478-017-0441-9
80. Nicolas, G. *et al.* SORL1 rare variants: a major risk factor for familial early-onset Alzheimer's disease. *Mol. Psychiatry* **21**, 831–836 (2015). doi: 10.1038/mp.2015.121
81. Gómez-Tortosa, E. *et al.* SORL1 Variants in Familial Alzheimer's Disease. *J. Alzheimer's Dis.* **61**, 1275–1281 (2018). doi: 10.3233/JAD-170590
82. Holstege, H. *et al.* Characterization of pathogenic SORL1 genetic variants for association with Alzheimer's disease: a clinical interpretation strategy. *Eur. J. Hum. Genet.* **25**, 973–981 (2017). doi: 10.1038/ejhg.2017.87
83. Beecham, G. W. *et al.* Genome-Wide Association Meta-analysis of Neuropathologic Features of Alzheimer's Disease and Related Dementias.

PLoS Genet. **10**, e1004606 (2014). doi: 10.1371/journal.pgen.1004606

84. Herold, C. *et al.* Family-based association analyses of imputed genotypes reveal genome-wide significant association of Alzheimer's disease with OSBPL6, PTPRG, and PDCL3. *Mol. Psychiatry* **21**, 1608–1612 (2016). doi: 10.1038/mp.2015.218
85. Andersson, C.-H. *et al.* A Genetic Variant of the Sortilin 1 Gene is Associated with Reduced Risk of Alzheimer's Disease. *J. Alzheimers. Dis.* **53**, 1353–1363 (2016). doi: 10.3233/JAD-160319
86. Philtjens, S. *et al.* Rare nonsynonymous variants in SORT1 are associated with increased risk for frontotemporal dementia. *Neurobiol. Aging* **66**, 181.e3-181.e10 (2018). doi: 10.1016/j.neurobiolaging.2018.02.011
87. Binzer, S. *et al.* Genetic analysis of the isolated Faroe Islands reveals *SORCS3* as a potential multiple sclerosis risk gene. *Mult. Scler. J.* **22**, 733–740 (2016). doi: 10.1177/1352458515602338
88. Scherzer, C. R. *et al.* Loss of Apolipoprotein E Receptor LR11 in Alzheimer Disease. *Arch. Neurol.* **61**, 1712–1716 (2004).doi: 10.1001/archneur.61.8.1200
89. Grear, K. E. *et al.* Expression of SORL1 and a novel SORL1 splice variant in normal and Alzheimers disease brain. *Mol. Neurodegener.* **4**, 46 (2009). doi: 10.1186/1750-1326-4-46
90. Caglayan, S. Identification of Alzheimer Disease Risk Genotype That Predicts Efficiency of SORL1 Expression in the Brain. *Arch. Neurol.* **69**, 373-379 (2012). doi: 10.1001/archneurol.2011.788
91. Felsky, D. *et al.* The SORL1 gene and convergent neural risk for Alzheimer's disease across the human lifespan. *Mol. Psychiatry* **19**, 1125–1132 (2014). doi: 10.1038/mp.2013.142
92. Karlsson Linnér, R. *et al.* Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. *Nat. Genet.* **51**, 245–257 (2019). doi: 10.1038/s41588-018-0309-3
93. Liu, M. *et al.* Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nat. Genet.* **51**, 237–244 (2019). doi: 10.1038/s41588-018-0307-5
94. Kichaev, G. *et al.* Leveraging Polygenic Functional Enrichment to Improve GWAS Power. *Am. J. Hum. Genet.* **104**, 65–75 (2019). doi: 10.1016/j.ajhg.2018.11.008

95. Duncan, L. *et al.* Significant Locus and Metabolic Genetic Correlations Revealed in Genome-Wide Association Study of Anorexia Nervosa. *Am J Psychiatry* **174**, 850–858 (2017). doi: 10.1176/appi.ajp.2017.16121402
96. Alemany, S. *et al.* New suggestive genetic loci and biological pathways for attention function in adult attention-deficit/hyperactivity disorder. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* **168**, 459–470 (2015). doi: 10.1002/ajmg.b.32341
97. Fabbri, C. & Serretti, A. Genetics of long-term treatment outcome in bipolar disorder. *Prog. Neuro-Psychopharmacology Biol. Psychiatry* **65**, 17–24 (2016). doi: 10.1016/j.pnpbp.2015.08.008
98. Okbay, A. *et al.* Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nat. Genet.* **48**, 624–633 (2016). doi: 10.1038/ng.3552
99. Baum, A. *et al.* A genome-wide association study implicates diacylglycerol kinase eta (DGKH) and several other genes in the etiology of bipolar disorder. *Mol. Psychiatry* **13**, 197–207 (2008). doi : 10.1038/sj.mp.4002012
100. Ollila, H. M. *et al.* Findings from bipolar disorder genome-wide association studies replicate in a Finnish bipolar family-cohort. *Mol. Psychiatry* **14**, 351–353 (2009). doi: 10.1038/mp.2008.122
101. Takata, A. *et al.* Nominal association between a polymorphism in DGKH and bipolar disorder detected in a meta-analysis of East Asian case-control samples. *Psychiatry Clin. Neurosci.* **65**, 280–285 (2011). doi: 10.1111/j.1440-1819.2011.02193.x
102. Howard, D. M. *et al.* Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nat. Neurosci.* **22**, 343–352 (2019). doi: 10.1038/s41593-018-0326-7
103. Howard, D. M. *et al.* Genome-wide association study of depression phenotypes in UK Biobank identifies variants in excitatory synaptic pathways. *Nat. Commun.* **9**, 1470 (2018). doi: 10.1038/s41467-018-03819-3
104. Wray, N. R. *et al.* Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat. Genet.* **50**, 668–681 (2018). doi: 10.1038/s41588-018-0090-3
105. Hyde, C. L. *et al.* Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nat. Genet.* **48**, 1031–1036 (2016). doi: 10.1038/ng.3623

106. Pan, L., Penney, J. & Tsai, L. Chromatin Regulation of DNA Damage Repair and Genome Integrity in the Central Nervous System. *J. Mol. Biol.* **426**, 3376–3388 (2014). doi: 10.1016/j.jmb.2014.08.001
107. Suberbielle, E. *et al.* Physiologic brain activity causes DNA double-strand breaks in neurons, with exacerbation by amyloid-β. *Nat. Neurosci.* **16**, 613–21 (2013). doi: 10.1038/nn.3356
108. Mehta, A. & Haber, J. E. Sources of DNA double-strand breaks and models of recombinational DNA repair. *Cold Spring Harb. Perspect. Biol.* **6**, a016428 (2014). doi: 10.1101/cshperspect.a016428
109. Rogakou, E. P., Pilch, D. R., Orr, A. H., Ivanova, V. S. & Bonner, W. M. DNA Double-stranded Breaks Induce Histone H2AX Phosphorylation on Serine 139. *J.Biol. Chem.* **273**, 5858-5868 (1998). doi : 10.1074/jbc.273.10.5858
110. Sedelnikova, O. A., Rogakou, E. P., Panyutin, I. G. & Bonner, W. M. Quantitative detection of (125)IdU-induced DNA double-strand breaks with gamma-H2AX antibody. *Radiat. Res.* **158**, 486–492 (2002). doi: 10.1667/0033-7587(2002)158[0486:qdoiid]2.0.co;2
111. Barral, S. *et al.* Phosphorylation of histone H2AX in the mouse brain from development to senescence. *Int. J. Mol. Sci.* **15**, 1554–1573 (2014). doi: 10.3390/ijms15011554
112. Firsanov, D. V, Solovjeva, L. V & Svetlova, M. P. H2AX phosphorylation at the sites of DNA double-strand breaks in cultivated mammalian cells and tissues. *Clin. Epigenetics* **2,** 283-297 (2011). doi:10.1007/s13148-011-0044-4
113. Bassing, C. H. & Alt, F. W. H2AX May Function as an Anchor to Hold Broken Chromosomal DNA Ends in Close Proximity. *Cell Cycle* **3**, 147–148 (2004). doi: 10.4161/cc.3.2.689
114. Nowak, E. *et al.* Radiation-induced H2AX phosphorylation and neural precursor apoptosis in the developing brain of mice. *Radiat. Res.* **165**, 155–164 (2006). doi : 10.1667/rr3496.1
115. Crowe, S. L., Movsesyan, V. A., Jorgensen, T. J. & Kondratyev, A. Rapid phosphorylation of histone H2A.X following ionotropic glutamate receptor activation. *Eur. J. Neurosci.* **23**, 2351–2361 (2006). doi: 10.1111/j.1460-9568.2006.04768.x
116. Crowe, S. L., Tsukerman, S., Gale, K., Jorgensen, T. J. & Kondratyev, A. D. Phosphorylation of Histone H2A.X as an Early Marker of Neuronal Endangerment following Seizures in the Adult Rat Brain. *J. Neurosci.* **31**, 7648–

7656 (2011). doi: 10.1523/JNEUROSCI.0092-11.2011

117. Fernando, R. N. *et al.* Cell cycle restriction by histone H2AX limits proliferation of adult neural stem cells. *Proc. Natl. Acad. Sci.* **108**, 5837–5842 (2011). doi: 10.1073/pnas.1014993108
118. Madabhushi, R. *et al.* Activity-Induced DNA Breaks Govern the Expression of Neuronal Early-Response Genes. *Cell* **161**, 1592-1605 (2015). doi:10.1016/j.cell.2015.05.032
119. Suberbielle, E. *et al.* DNA repair factor BRCA1 depletion occurs in Alzheimer brains and impairs cognitive function in mice. *Nat. Commun.* **6**, 8897 (2015). doi: 10.1038/ncomms9897
120. Jacobsen, E., Beach, T., Shen, Y., Li, R. & Chang, Y. Deficiency of the Mre11 DNA repair complex in Alzheimer's disease brains. *Mol. Brain Res.* **128**, 1–7 (2004). doi : 10.1016/j.molbrainres.2004.05.023
121. Shen, X., Chen, J., Li, J., Kofler, J. & Herrup, K. Neurons in Vulnerable Regions of the Alzheimer's Disease Brain Display Reduced ATM Signaling. *eNeuro* **3**, ENEURO.0124-15.2016 (2016). doi: 10.1523/ENEURO.0124-15.2016
122. Sedelnikova, O. A. *et al.* Senescing human cells and ageing mice accumulate DNA lesions with unrepairable double-strand breaks. *Nat. Cell Biol.* **6**, 168–170 (2004). doi : 10.1038/ncb1095
123. Khan, F. A. & Ali, S. O. Physiological Roles of DNA Double-Strand Breaks. *J. Nucleic Acids* **2017**, 1–20 (2017). doi: 10.1155/2017/6439169
124. Mullaart, E., Boerrigter, M. E. T. I., Ravid, R., Swaab, D. F. & Vijg, J. Increased levels of DNA breaks in cerebral cortex of Alzheimer's disease patients. *Neurobiol. Aging* **11**, 169–173 (1990). doi: 10.1016/0197-4580(90)90542-8
125. Adamec, E., Vonsattel, J. P. & Nixon, R. A. DNA strand breaks in Alzheimer's disease. *Brain Res.* **849**, 67–77 (1999). doi: 10.1016/s0006-8993(99)02004-1
126. Simpson, J. E. *et al.* Population variation in oxidative stress and astrocyte DNA damage in relation to Alzheimer-type pathology in the ageing brain. *Neuropathol. Appl. Neurobiol.* **36**, 25–40 (2010). doi: 10.1111/j.1365-2990.2009.01030.x
127. Simpson, J. E. *et al.* A neuronal DNA damage response is detected at the earliest stages of Alzheimer's neuropathology and correlates with cognitive impairment in the Medical Research Council's Cognitive Function and Ageing Study ageing brain cohort. *Neuropathol. Appl. Neurobiol.* **41**, 483–496 (2015). doi: 10.1111/nan.12202

128. Shanbhag, N. M. *et al.* Early neuronal accumulation of DNA double strand breaks in Alzheimer's disease. *Acta Neuropathol. Commun.* **7**, 77 (2019). doi: 10.1186/s40478-019-0723-5
129. Raza, M. U. *et al.* DNA Damage in Major Psychiatric Diseases. *Neurotox. Res.* **30**, 251-267 (2016). doi: 10.1007/s12640-016-9621-9
130. Lovell, M. A. & Markesbery, W. R. Oxidative damage in mild cognitive impairment and early Alzheimer's disease. *J. Neurosci. Res.* **85**, 3036–3040 (2007). doi : 10.1002/jnr.21346
131. Sharma, V. *et al.* Oxidative stress at low levels can induce clustered DNA lesions leading to NHEJ mediated mutations. *Oncotarget* **7**, 25377–25390 (2016). doi: 10.18632/oncotarget.8298
132. Katsube, T. *et al.* Most hydrogen peroxide-induced histone H2AX phosphorylation is mediated by ATR and is not dependent on DNA double-strand breaks. *J. Biochem.* **156**, 85–95 (2014). doi: 10.1093/jb/mvu021
133. Scholz, D. *et al.* Rapid, complete and large-scale generation of post-mitotic neurons from the human LUHMES cell line. *J. Neurochem.* **119**, 957–971 (2011). doi: 10.1111/j.1471-4159.2011.07255.x
134. Scholz, D., Chernyshova, Y. & Leist, M. Control of Aβ release from human neurons by differentiation status and RET signaling. *Neurobiol. Aging* **34**, 184–199 (2013). doi: 10.1016/j.neurobiolaging.2012.03.012
135. Austria, B. *et al.* Efficient and versatile CRISPR engineering of human neurons in culture to model neurological disorders. *Wellcome Open Res* **1**, 13 (2016). doi:10.12688/wellcomeopenres.10011.1
136. Mertens, J., Reid, D., Lau, S., Kim, Y. & Gage, F. H. Aging in a Dish: iPSC-Derived and Directly Induced Neurons for Studying Brain Aging and Age-Related Neurodegenerative Diseases. *Annu. Rev. Genet.* **52**, 271–293 (2018). doi: 10.1146/annurev-genet-120417-031534
137. De Marco, M. & Venneri, A. Volume and Connectivity of the Ventral Tegmental Area are Linked to Neurocognitive Signatures of Alzheimer's Disease in Humans. *J. Alzheimer's Dis.* **63**, 167–180 (2018). doi: 10.3233/JAD-171018
138. Morris, J. C., Drazner, M., Fulling, K., Grant, E. A. & Goldring, J. Clinical and Pathological Aspects of Parkinsonism in Alzheimer's Disease. *Arch. Neurol.* **46**, 651-657 (1989).doi : 10.1001/archneur.1989.00520420071025
139. Perez, S. E. *et al.* Nigrostriatal Dysfunction in Familial Alzheimer's Disease-Linked APPswe/PS1E9 Transgenic Mice. *J. Neurosci.* **25**, 10220-10229

(2005). doi:10.1523/JNEUROSCI.2773-05.2005

140. Nobili, A. *et al.* Dopamine neuronal loss contributes to memory and reward dysfunction in a model of Alzheimer's disease. *Nat. Commun.* **8**, 14727 (2017). doi: 10.1038/ncomms14727

141. Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013). doi: 10.1038/nprot.2013.143

142. Giau, V. Van, Lee, H., Shim, K. H., Bagyinszky, E. & An, S. S. A. Genome-editing applications of CRISPR–Cas9 to promote in vitro studies of Alzheimer's disease. *Clin. Interv. Aging* **13**, 221–233 (2018). doi: 10.2147/CIA.S155145

143. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012). doi: 10.1038/nmeth.2019

144. Alt, F. W. & Schwer, B. DNA double-strand breaks as drivers of neural genomic change, function, and disease. *DNA Repair (Amst).* **71**, 158–163 (2018). doi: 10.1016/j.dnarep.2018.08.019

145. Simpson, J. E. *et al.* A neuronal DNA damage response is detected at the earliest stages of Alzheimer's neuropathology and correlates with cognitive impairment in the Medical Research Council's Cognitive Function and Ageing Study ageing brain cohort. *Neuropathol. Appl. Neurobiol.* **41**, 483–496 (2015). doi: 10.1111/nan.12202

146. Mitra, J. *et al.* Motor neuron disease-associated loss of nuclear TDP-43 is linked to DNA double-strand break repair defects. *Proc. Natl. Acad. Sci.* **116**, 4696–4705 (2019). doi: 10.1073/pnas.1818415116.

147. Siddiqui, M. S. *et al.* γH2AX is increased in peripheral blood lymphocytes of Alzheimer's disease patients in the South Australian Neurodegeneration, Nutrition and DNA Damage (SAND) study of aging. *Mutat. Res. Toxicol. Environ. Mutagen.* **829**–**830**, 6–18 (2018). doi: 10.1016/j.mrgentox.2018.03.001

148. Reitz, C. *et al.* Independent and epistatic effects of variants in VPS10-d receptors on Alzheimer disease risk and processing of the amyloid precursor protein (APP). *Transl. Psychiatry* **3**, e256 (2013). doi: 10.1038/tp.2013.13

149. Blanco-Rodríguez, J. γH2AX marks the main events of the spermatogenic process. *Microsc. Res. Tech.* **72**, 823–832 (2009). doi: 10.1002/jemt.20730

150. Noguchi, K., Gel, Y. R., Brunner, E. & Konietschke, F. **nparLD** : An *R* Software Package for the Nonparametric Analysis of Longitudinal Data in Factorial Experiments. *J. Stat. Softw.* **50**, 1–23 (2012). doi: 10.18637/jss.v050.i12

151. Alvarez-Castelao, B. *et al.* Cell-type-specific metabolic labeling of nascent proteomes in vivo. *Nat. Biotechnol.* **35**, 1196–1201 (2017). doi: 10.1038/nbt.4016
152. Ruan, C.-S. *et al.* Sortilin inhibits amyloid pathology by regulating non-specific degradation of APP. *Exp. Neurol.* **299**, 75–85 (2018). doi: 10.1016/j.expneurol.2017
153. Massey, P. V. *et al.* Differential Roles of NR2A and NR2B-Containing NMDA Receptors in Cortical Long-Term Potentiation and Long-Term Depression. *J. Neurosci.* **24**, 7821–7828 (2004). doi : 10.1523/JNEUROSCI.1697-04.2004
154. Ismail, I. H. & Hendzel, M. J. The γ-H2A.X: Is it just a surrogate marker of double-strand breaks or much more? *Environ. Mol. Mutagen.* **49**, 73–82 (2008). doi: 10.1002/em.20358
155. Kamat, P. K. *et al.* Mechanism of Oxidative Stress and Synapse Dysfunction in the Pathogenesis of Alzheimer's Disease: Understanding the Therapeutics Strategies. *Mol. Neurobiol.* **53**, 648–661 (2016). doi: 10.1007/s12035-014-9053-6
156. Butterfield, D. A., Swomley, A. M. & Sultana, R. Amyloid β-peptide (1-42)-induced oxidative stress in Alzheimer disease: importance in disease pathogenesis and progression. *Antioxid. Redox Signal.* **19**, 823–35 (2013). doi: 10.1089/ars.2012.5027
157. Zhang, Y., Li, P., Feng, J. & Wu, M. Dysfunction of NMDA receptors in Alzheimer's disease. *Neurol. Sci.* **37**, 1039–1047 (2016). doi: 10.1007/s10072-016-2546-5
158. Bordji, K., Becerril-Ortega, J., Nicole, O. & Buisson, A. Activation of Extrasynaptic, But Not Synaptic, NMDA Receptors Modifies Amyloid Precursor Protein Expression Pattern and Increases Amyloid- Production. *J. Neurosci.* **30**, 15927–15942 (2010). doi: 10.1523/JNEUROSCI.3021-10.2010
159. Schildknecht, S. *et al.* Generation of genetically-modified human differentiated cells for toxicological tests and the study of neurodegenerative diseases. *ALTEX* **30**, 427–444 (2013). doi: 10.14573/altex.2013.4.427
160. Greenbaum, D., Colangelo, C., Williams, K. & Gerstein, M. Comparing protein abundance and mRNA expression levels on a genomic scale. *Genome Biol.* **4**, 117 (2003).doi : 10.1186/gb-2003-4-9-117
161. Pelley J. Protein Synthesis and Degradation. *Elsevier's Integrated Biochemistry* 147–158 (2007). doi:10.1016/B978-0-323-03410-4.50023-7

162. Montecucco, A., Zanetta, F. & Biamonti, G. Molecular mechanisms of etoposide *EXCLI J.* **14**, 95–108 (2015). doi: 10.17179/excli2015-561
163. Austin, C. A. & Marsh, K. L. Eukaryotic DNA topoisomerase IIβ. *BioEssays* **20**, 215–226 (1998). doi: 10.1002/(SICI)1521-1878(199803)20:3<215::AID-BIES5>3.0.CO;2-Q
164. Rehmani, N., Zafar, A., Arif, H., Hadi, S. M. & Wani, A. A. Copper-mediated DNA damage by the neurotransmitter dopamine and L-DOPA: A pro-oxidant mechanism. *Toxicol. Vitr.* **40**, 336–346 (2017). doi: 10.1016/j.tiv.2017.01.020
165. Fazeli, G., Oli, R. G., Schupp, N. & Stopper, H. The Role of the Dopamine Transporter in Dopamine-Induced DNA Damage. *Brain Pathol.* **21**, 237–248 (2011). doi: 10.1111/j.1750-3639.2010.00440.x
166. Anderson, L., Henderson, C. & Adachi, Y. Phosphorylation and rapid relocalization of 53BP1 to nuclear foci upon DNA damage. *Mol. Cell. Biol.* **21**, 1719–1729 (2001).doi: 10.1128/MCB.21.5.1719-1729.2001
167. Ward, I. M., Minn, K., Jorda, K. G. & Chen, J. Accumulation of checkpoint protein 53BP1 at DNA breaks involves its binding to phosphorylated histone H2AX. *J. Biol. Chem.* **278**, 19579–82 (2003). doi: 10.1074/jbc.C300117200
168. Tsutsui, K., Tsutsui, K., Sano, K., Kikuchi, A. & Tokunaga, A. Involvement of DNA topoisomerase IIbeta in neuronal differentiation. *J. Biol. Chem.* **276**, 5769–78 (2001). doi: 10.1074/jbc.M008517200
169. Madabhushi, R., Pan, L. & Tsai, L.-H. DNA damage and its links to neurodegeneration. *Neuron* **83**, 266–282 (2014). doi: 10.1016/j.neuron.2014.06.034
170. Qiu, W. Q. *et al.* Degradation of amyloid beta-protein by a serine protease-alpha2-macroglobulin complex. *J. Biol. Chem.* **271**, 8443–8451 (1996). doi: 10.1074/jbc.271.14.8443
171. Paladini, C. A., Robinson, S., Morikawa, H., Williams, J. T. & Palmiter, R. D. Dopamine controls the firing pattern of dopamine neurons via a network feedback mechanism. *Proc. Natl. Acad. Sci.* **100**, 2866–2871 (2003). doi: 10.1073/pnas.0138018100
172. Müller, U. C., Deller, T. & Korte, M. Not just amyloid: physiological functions of the amyloid precursor protein family. *Nat. Rev. Neurosci.* **18**, 281–298 (2017). doi: 10.1038/nrn.2017.29
173. Vardarajan, B. N. *et al.* Coding mutations in SORL1 and Alzheimer's disease The Taub Institute for Research on Alzheimer's Disease and the Aging Brain

HHS Public Access. *Ann Neurol* **77**, 215–227 (2015). doi: 10.1002/ana.24305

174. Bellenguez, C. *et al.* Contribution to Alzheimer's disease risk of rare variants in TREM2, SORL1, and ABCA7 in 1779 cases and 1273 controls. *Neurobiol. Aging* **59**, 220.e1-220.e9 (2017). doi: 10.1016/j.neurobiolaging.2017.07.001
175. Thonberg, H. *et al.* Identification and description of three families with familial Alzheimer disease that segregate variants in the SORL1 gene. *Acta Neuropathol. Commun.* **5**, 43 (2017). doi: 10.1186/s40478-017-0441-9
176. Gómez-Tortosa, E. *et al.* SORL1 Variants in Familial Alzheimer's Disease. *J. Alzheimer's Dis.* **61**, 1275–1281 (2018). doi: 10.3233/JAD-170590
177. Caglayan, S. *et al.* Identification of Alzheimer Disease Risk Genotype That Predicts Efficiency of SORL1Expression in the Brain. *Arch. Neurol.* **69**, 373 (2012). doi: 10.1001/archneurol.2011.788
178. Kitago, Y. *et al.* Structural basis for amyloidogenic peptide recognition by sorLA. *Nat. Struct. Mol. Biol.* **22**, 199-206 (2015). doi:10.1038/nsmb.2954
179. Inui, M. *et al.* Rapid generation of mouse models with defined point mutations by the CRISPR/Cas9 system. *Sci. Rep.* **4**, 5396 (2015). doi: 10.1038/srep05396
180. Paquet, D. *et al.* Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature* **533**, 125–129 (2016). doi: 10.1038/nature17664
181. Richardson, C. D., Ray, G. J., Dewitt, M. A., Curie, G. L. & Corn, J. E. Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat. Biotechnol.* **34**, 339-344 (2016). doi: 10.1038/nbt.3481
182. Lin, S., Staahl, B. T., Alla, R. K. & Doudna, J. A. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *Elife* **3**, e04766 (2014). doi: 10.7554/eLife.04766
183. Yang, L. *et al.* Optimization of scarless human stem cell genome editing.*Nucleic Acids Res.* **41**, 9049-9061 (2013). doi:10.1093/nar/gkt555
184. Quax, T. E. F., Claassens, N. J., Söll, D. & Oost, J. van der. Codon Bias as a Means to Fine-Tune Gene Expression. *Mol. Cell* **59**, 149-161 (2015). doi: 10.1016/j.molcel.2015.05.035
185. Anna, A. & Monika, G. Splicing mutations in human genetic disorders: examples, detection, and confirmation. *J.Appl. Genet.* **59**, 253-268 (2018). doi:10.1007/s13353-018-0444-7

186. Kosicki, M., Tomberg, K. & Bradley, A. Repair of double-strand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements. *Nat. Biotechnol.* **36**, 765-771 (2018). doi: 10.1038/nbt.4192

187. MaGarridota-, J., Casafont, I., Tapia, O., Berciano, M. T. & Lafarga, M. Neuronal accumulation of unrepaired DNA in a novel specific chromatin domain: structural, molecular and transcriptional characterization. *Acta Neuropathol. Commun.* **4**, 41 (2016). doi:10.1186/s40478-016-0312-9

188. Mariya H. Quantitative nature of overexpression experiments. *Mol.Biol.Cell.* **26**, 3932-3939 (2015). doi: 10.1091/mbc.E15-07-0512